Ways of Mapping the World

Maths and cartography

OCEANO ATLANTICO

Everything is mathematical

Ways of Mapping the World

Ways of Mapping the World

Maths and cartography

Raúl Ibáñez

Everything is mathematical

To the memory of my father; I miss him so much. To my mother, a strong, resolute woman. To both of them I owe my life and much more.

To my wife Ana, and my children Aitor and Vanessa, who are my whole life

© 2010 Raúl Ibáñez (text)
© 2013, RBA Contenidos Editoriales y Audiovisuales, S.A.U.
Published by RBA Coleccionables, S.A.
c/o Hothouse Developments Ltd
91 Brick Lane, London, E1 6QL

Localisation: Windmill Books Ltd.
Photographs: Getty Images

All rights reserved. No part of this publication can be reproduced, sold or transmitted by any means without permission of the publisher.

ISSN: 2050-649X

Printed in Spain

Contents

Preface 7

Chapter 1. The Shape of the Earth 11
Our planet: is it flat or is it round? 13
Direct evidence of the Earth's sphericity 16
Mediaeval thinking 19
From the ellipsoid model to the geoid 22

Chapter 2. The Dimensions of the Earth 25
The estimates of Eudoxus and Archimedes 25
Eratosthenes' measurement 26
Posidonius' measurement and Colombus' error 31
The triangulation technique 32

Chapter 3. Meridians, Parallels and Great Circles 35
Latitude and the parallels 35
Longitude and the meridians 40
The problem of the determination of longitude 42
Great circles, the geodesics of the sphere 46
The curvature of the great circles 49

Chapter 4. In Search of an Accurate Map of the Earth 51
What does it mean to say that a map is accurate? 52
A double problem: scale and cartographic projection 54
A projection that preserves distances also preserves the shortest paths 58
If a projection preserves distances it is equivalent to its preserving the length of the curves 60
A projection that preserves distances also preserves bearings 62
A projection that preserves distances also preserves areas 64
In search of isometric projection 65

Chapter 5. The Archimedes Projection or Lambert Equal-Area Cylindrical Projection 71
Definition and cartographic properties 72

Cylindrical and pseudocylindrical projections 76
The use of equal-area maps 81

Chapter 6. The Central or Gnomonic Projection 87
Definition and cartographic properties 89
Azimuthal projections 95
The use of maps created by gnomonic projection 100

Chapter 7. Stereographic Projection 105
Definition and cartographic properties 107
The use of maps created by stereographic projection 112
Conic projections 123
The Lambert conformal conic projection 127

Chapter 8. What Euler Told the Cartographer 129
Equal-area conformal projections 130
Do accurate maps of the Earth exist? 131
The Gauss curve, a return to the problem of cartography 135
The terrestrial globe 140
Equidistant projections 144
  Equidistant cylindrical projection 144
  Equidistant azimuthal projection 145
  Equidistant conic projection 148

Chapter 9. The Mercator Projection 149
Definition and cartographic properties 149
The transverse Mercator projection 160
The oblique Mercator projection 162
Peters versus Mercator 164

Epilogue 171

Bibliography 173

Index 174

Preface

This book's main purpose is to discuss the geometry of maps. First, we should ask ourselves what a map is. According to the dictionary, a map is the "diagrammatic representation on a flat surface of a part or the whole of the surface of the Earth or of other planets, the sky, heavenly bodies etc."

However, the reader will agree with me when I say that we don't need to pick up a dictionary to know what a map is. Maps are familiar objects for all of us, they are part of daily life and we come across them all the time from morning to night, both at work and in private life, whether actively or passively. We all remember, perhaps with a certain nostalgia, the maps that decorated the walls of our childhood schoolrooms. When we read the newspaper or see the news on television we are met with an infinity of maps. We might find, for example, a map of the world with information on religions, languages or populations, a map showing different levels of pollution, of the production of certain goods or the road accident rate for a certain territory, or a map explaining the economic situation of different countries in the European Union, or one showing some of the world's war zones. And then there is, of course, the map we most look out for – the weather map. Watching any TV documentary, whether it be on nature, history, geography, economics or any other issue, or reading any specialised or popular culture book or magazine, we find explanatory maps that help us to assimilate the information and to place the various data in the right context.

We also find maps in films and science-fiction books, imaginary maps such as those in *The Lord of the Rings* or *Treasure Island,* real maps such as in adventure and war films, *Casablanca* or *War Games*, or fun maps like the one in the cartoon film *The Emperor's New Groove*, in which the protagonists actually walk on the map. We can also find them in art, from the meaningful maps appearing in the work of the Dutch baroque painter Johannes Vermeer to the Dymaxion map by the contemporary US painter Jasper Johns, or the maps of the world by the Italian artist Alighiero Boetti.

When planning our holidays we get hold of different maps to study the different choices for our journeys, to plan them, and finally, to follow the routes they show. Similarly, when we make a car journey we need a road map, and if we take a look round an unfamiliar town or city then we need a street map so as not to get lost. When walking around the town or city where we live we can see maps on company advertisements, in the windows of travel agents, in children's clothes shops, or in bookshops announcing the start of the school year. We also find maps on some companies' logos and on their products' wrappings.

Maps can also form part of our work. A large number of jobs involve, to a greater or lesser extent, the use of maps: maps for marine or air navigation, political maps, urban maps, maps for communications (railways, roads, etc.), topographical maps, morphological maps, different types of scientific maps (botanical, geological, climatic, geographical, oceanographic, seismological and so on) economic and statistical maps, artistic maps used for advertisements or for tourism, land registry maps showing plots belonging to different owners, and many, many more. Thus we can say we know what maps are, that we know all about those objects that we handle daily for very diverse reasons in our day-to-day lives.

On the other hand, the world map shown below (which is the map drawn up with the Mercator projection that we shall study in Chapter 9) is the map with which we are most familiar. We - at least those of a certain generation - have seen this map since our childhood as a representation of the Earth and so, therefore, we subconsciously accept it as an unalterable reality. We all know, this is a good map, an accurate map or, as someone told me one day: "This is the true map." Let's take another look at the world map and pose some simple questions: what route should we take to go from Rome (or, if we want to stretch it out a bit more, from Baku) to Washington D.C.? As the shortest path between two points is a straight line, the answer would seem to be the 40° N parallel. However, as will be seen in Chapter 3, on a sphere the shortest path between any two points is the great circle that passes through those points, and in this case its image on the plane is not parallel 40 ° N. That is one of the reasons why flights from Rome to Washington do not follow parallel 40 ° N, as one might

expect, but instead go up northwards and then come down southwards following the great circle. (If we had opted for Baku then the route to Washington would take us almost to the North Pole.) We can conclude that for our world map, the shortest paths have not been preserved.

As we all know, it is normal practice to show the scale at the bottom of the map. Given two places on the Earth, what is the distance between them? In this case we take a ruler and measure the distance between the places and afterwards convert that measurement into the distance by using the map's scale. However, as we said before, on the plane we should measure the length of the image curve of the great circle and not the length of a straight line between the two points. But even so, the result we would get would still be incorrect, and this is due to the fact that our map does not preserve the lengths of the curves, nor the distances and, in reality, the scale cannot be said to be uniform throughout the map. Following on with this line of thinking, let's look at whether the area is preserved in the Mercator map. It is well known that on this map Greenland is shown as being too big; it appears to be as large as Africa. The reality, however, is that Greenland has an area of approximately 2,175,600 km^2, and Africa an area of some 29,800,000 km^2; the area, therefore, has been greatly distorted. Finally, let's look at whether maps preserve the directions and, in short, the angles. The angle between meridians and parallels is 90°, and it is on our map too. However, if we take a look at the map on the next page, which is the map we get if we look at the Earth from infinite space (this type of classical projection is called an orthographic projection), here it is not true, in other words, the angles are not preserved on it.

Maps, then, do not follow the behaviour we expect regarding matters such as distances, shortest paths, areas and directions, so perhaps we don't know so much about them after all. What's more, not only are there the two types of map we have just mentioned, but a great number of different kinds of planar representations of the Earth: the Mercator, the orthographic, Lambert's cylindrical equal-area, Alber's conic equal-area, the Mollweide, the Gall-Peters orthographical, the Eckert IV, the central, the stereographical, the bipolar oblique conformal conic, the equidistant cylindrical, the equidistant azimuthal, the Winkel-Tripel, the Van der Gritten, the UTM, the Bonne, the Eckert I-IV, Goodes' homolosine, the Hammer, the Werner, the Briesemeister, the Behrmann cylindrical equal-area, the Robinson projection and many others. In fact, in his book *Flattening the Earth,* the cartographer John Snyder mentions 300 different cartographic projections. So, we might ask, why are there

so many types of maps? Which is the correct one? Why aren't the others correct? How should we accurately draw a map of the Earth? What does 'accurately' mean?

This book's goal is, partly, to answer those questions as well as to provide greater knowledge of these ubiquitous objects. The key tool required to dig deeper into the study of such questions is differential geometry, as we can see from looking at the curricula on cartography in different university degree courses, or by looking through the usual literature on cartography. However, our objective in this book is to avoid technical terminology and to be fundamentally intuitive – 'geometrical' – and so the tools that we shall use are the ones found in classical geometry (basically Euclidean geometry and trigonometry). The approximations that appear in many of the arguments disappear if we take them to the limit, but let's note that in that case we would be doing nothing more than simply introducing differential and integral calculus, and then we would find ourselves within differential geometry.

Chapter 1

The Shape of the Earth

> *"Firstly," replied Socrates, "I'm convinced that if the Earth is in the middle of the sky and is spherical, it has no need of the air or any other support to stop it from falling…*
> *What's more, I'm sure that the Earth is very big and we don't occupy more than this part that stretches from Phasis to the Pillars of Hercules, spread around the sea like ants and frogs around a lake…*
> *It is said, dear Simmias, that if you look at this Earth from a high point, it looks like one of those leather balls…"*
>
> Plato, *Phaedo*, or *The Immortality of the Soul*
> (4th century BC)

Before we start to draw up and study maps of the planet on which we live, we have to concentrate on its shape and dimensions by determining the position of the points on its surface or by studying some of its outstanding geometric features, just as has been done since ancient times. Even back then, Claudius Ptolemy said in his *Geography* that: "…The first thing that has to be researched [to map the world] is the shape, the size and the position of the Earth in respect to its surroundings [the sky], so it will be possible to discuss the known part, what it is like and how big it is […]. Such things belong to the most noble and beautiful of intellectual adventures, that is, to demonstrate human intelligence through mathematics… [the nature of] the Earth through its portrait…".

That is precisely the aim of geodesics, a word from the Greek *geo* ('earth') and *daio* ('divide'), so it means 'to divide the Earth', and is the science that studies the shape and dimensions of the planet, its field of gravity and its temporal variations. Among the basic tools of geodesics are the different branches of mathematics and in particular, geometry, whose etymological relationship with geodesics shows the

CLAUDIUS PTOLEMY (c. 90–c. 170)

Little is known about the life of this astronomer, mathematician and geographer except that he was a Roman astronomer of Greek or Egyptian extraction, who lived and worked in Alexandria. He was the author of two treatises that had great influence on Islamic and European science, the *Almagest* (*The Great Treatise*, originally *Mathematical Treatise*) and *Geography*. *Almagest* is a treatise in which Ptolemy, very much influenced by Hipparchus, covers and expands on Greek astronomical knowledge, as well as covering the mathematics needed for studying it. In the work, he provides a detailed exposition of the mathematical theory that describes the movement of the Sun, the Moon and the planets. His model is geocentric and describes the movement of those heavenly bodies, which he considers to be spherical, by making use of combined circular movements (epicycles). The *Almagest* also contains a list of the stars. Ptolemy also propagated his ideas in a simpler style in his work *Planetary Hypotheses*. *Geography* is a re-compilation of the geographical knowledge of the world in his times. It develops and gives instructions on how to create maps of the known world (the Ecumene) and of the Roman provinces by making use of a grid of coordinates. Ptolemy's maps (the surviving reproductions of which are a product of the first printed editions of the 15th century) had the great virtue of having been created by using geometric projections. However, they were very imprecise due to the erroneous data for the territory beyond the Roman Empire, and even within the empire, and for the underestimation that he made of the size of the Earth. In his works *Analemma* and *Planispherium* the orthographical and stereographic projections are both explained. He also wrote *Harmonics*, on music, *Optics* and *Tetrabiblos*, a treatise on astrology.

A reconstruction of one of the world maps in Ptolemy's Geography, which appears in the Cosmographia *by Johannes de Armsshein (or Nicolas Germanus) from 1482.*

important link between the two: the word geometry derives from the Greek *geo* ('earth') and *metria* ('measure'), that is, 'to measure the Earth'.

We'll start off a first section of themes covering the first three chapters dedicated to the study of the Earth, 'that dark object of desire', its shape and size, the geographic coordinates and the maximum circles.

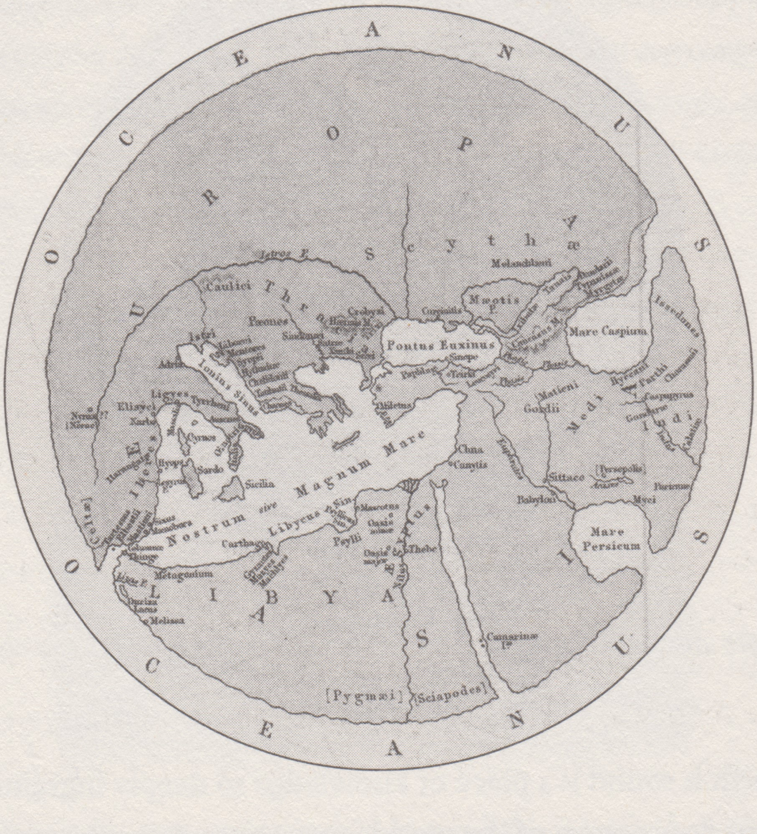
Our planet: is it flat or is it round?

Asking the question of what shape the Earth is may nowadays seem childish and even a little patronising, as we've all known very well since our school days that our planet is as round as a ball (mathematically, a sphere), though a little flattened at the poles (more like what we mathematicians call an ellipsoid). Also from our school days stems the belief that the human race was convinced that the world was flat until Christopher Columbus put them right by proving to his contemporaries that it was round.

Satellite images of the Earth show us that our planet is round and not flat.

That the Earth is round is a piece of knowledge so deeply ingrained within us, in our thinking, our conception of the world, our society, that we are never, ever going to question it. But what direct evidence have we that the Earth is really round? We have seen a great many images taken from satellites in which it can be seen that our world is a huge sphere. However, discounting the faint possibility of a conspiracy and that the images aren't authentic, that they've been manipulated, it is not a fact that we can personally verify, as the Greek philosopher Aristotle (384–322 BC) put it in his treatise *On the Heavens*, "through phenomena accessible to our senses".

Since ancient times, in fact, we inhabitants of this planet have sensed, and even believed, that we are living on a flat world, which except for the reliefs caused by mountains, valleys and oceans, is planar. And, in fact, that is what was believed up to about 2,500 years ago by the different cultures, such as the Egyptians, the Babylonians, the Chinese and even the Greeks. The first Greek references to the shape of the Earth are found in the work of the Greek poet Homer (9th century BC), which contains the geographic and cosmological knowledge of his times. The Greeks thought that the Earth was a flattened disc, held up by the air, containing the world known at that time and surrounded by water, the great river-ocean whose waters dashed against its edges. That was the thinking of the Ionian school of philosophy and, in particular, of Anaximander (c. 610–c. 546 BC), a follower of Thales of Miletus, who was the author of the first known world map.


OCEANUS
OCEANUS
EUROPA
Scythæ
Pontus Euxinus
Mare Caspium
Nostrum sive Magnum Mare
LIBYA
ASIA
Mare Persicum
Indi
Medi
Babylon
Sicilia
Sardo
Carthago
Thebe
[Pygmæi]
[Sciapodes]

A reconstruction of the Hecataeus map, based on Anaximander's map, which is the first known representation of the Ecumene, i.e., the known world.

THE EARTH IN COSMOLOGICAL MYTHOLOGY.

All the peoples of the ancient world (the Babylonians, Egyptians, Chinese, Greeks, indigenous Americans and so on) had their own myths about the origins and structure of the world. They imagined the Earth as more or less flat, whether resting on an ocean of water, floating in the air, or on some mythological animal.

For the Babylonians, the Earth was a flat disc that floated on an ocean and was covered by the heavenly canopy, a semi-spherical metallic hood from which the stars hung. Over the heavenly canopy there was more water, the superior water, which at times filtered through, creating rain. In some myths, for example African ones, the Earth rests on a serpent swimming in the ocean, while for Hindus it rests on four elephants situated on a giant tortoise which is also swimming in the sea. For the Egyptians and the Chinese, the flat Earth was rectangular and also floated on water, while the heavenly canopy rested on two mountain chains or on four mountains located at the corners of the world.

For the Mayans, as well as for other American cultures, the world was a flat rectangular sheet underneath the sky, formed by thirteen superimposed horizontal planes and forming a pyramidal structure with the chief god at the top. Under this Earth is the infraworld, composed of nine flat horizontal layers like an inverted pyramid. The idea of a vertical cosmology formed by parallel flat worlds is also common in Buddhism.

The idea of the Earth being spherical is normally attributed to the Greek philosopher and mathematician Pythagoras (c. 570–c. 500 BC), though there is some controversy over the issue. We don't know if this assertion is based on physical observations or simply on reasoning of a more philosophical nature, as a sphere is the most perfect geometric form and therefore this planet, the home of the human race and the work of the gods, also had to have this shape. The idea that the world is round is mentioned by Plato in his dialogue *Phaedo, or The Immortality of the Soul*. But in particular this reasoning is included by Aristotle in his treatise *On the Heavens*, in which he provides physical and logical arguments to support the theory, as well as giving the first known estimation of the radius of the Earth: "All the mathematicians who attempt to calculate the size of the Earth's circumference say that is it 400,000 stadia." The question of the dimensions of the globe is discussed in the next chapter.

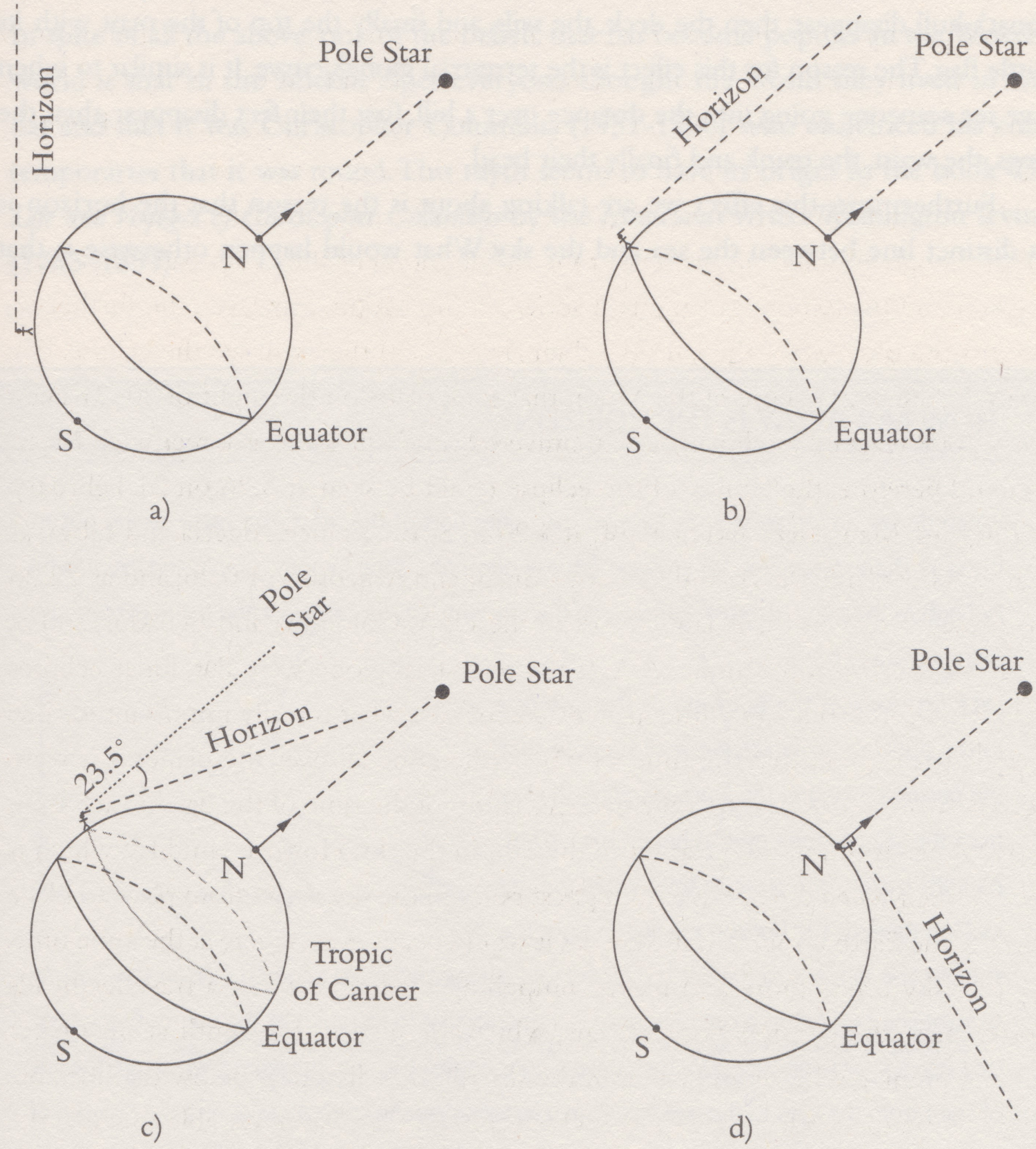
Direct evidence of the Earth's sphericity

By making use of the fact that the arguments Aristotle provided to prove the Earth was spherical are still valid today, we shall answer the question asked at the beginning of this section about what direct evidence can prove to us that the Earth really is round. By looking up at the sky, just as the Greeks did, we can obtain the first piece of evidence by observing that the heavenly bodies that we see moving about up there – the Moon, the Sun and the planets – are also round. Furthermore, if we look at the shadow of the Earth on the Moon during a lunar eclipse, when the Earth is situated between the Sun and the Moon, that shadow is always circular.

Though not quite so obvious, lunar eclipses offer more evidence. These are occur at a single moment in universal time, at the moment the Earth passes between the Sun and the Moon, casting its shadow on the Moon. However, the further to the east the observer is, the later (by their own clock) they will see the eclipse. For example, the total eclipse of the Moon that took place on the night of 20–21 February 2008 reached its climax at 3:26 universal time (the same as Greenwich Mean Time). Therefore, the climax of the eclipse could be seen at 3.26 on 21 February in England, Mauritania and Senegal; at 4.26 in Spain, France, Algeria and Libya; at 0.26 in Greenland, the Atlantic coast of Brazil and Argentina at 0.26; and at 22.26 the previous day on the Atlantic coast of the USA, Colombia and Ecuador; and at 21.26 in Mexico and central USA. If the terrestrial globe were flat, lunar eclipses would be seen at the same time in all places. In fact, what is really important for our purposes is to note that the time would be the same all over the planet if it were flat. The reason for this is that the determination of the time of the day in each time zone has its origin in the position of the Sun in the sky. However, midday, which is the moment when the Sun is at its highest point in the sky, varies from place to place because the Earth is round. If it were flat it would occur everywhere at the same time.

The sky offers more and more compelling evidence: when a traveller heads northwards, the constellations of stars, which are seen to the south at his or her starting point, get lower and lower in the sky till they disappear below the horizon. At the same time, other stars appear which the traveller was never able to see at the place where the journey began, and these stars will rise higher and higher in the sky. More specifically, if we are in the southern hemisphere we cannot see Polaris (the Pole Star); on moving northwards it will appear on the horizon when we pass over the Equator and little by little it will rise up in the sky until it is right over our heads if we go as far as the North Pole itself. That would not happen in a flat world, in which we would see the same constellations from everywhere.

If we lower our sights a little and study the horizon, we shall also find some evidence of the roundness of our planet. To start with, let's place ourselves on the coast facing the open sea, or on the deck of a ship on the high seas, and we'll find that the line of the horizon curves down quite a lot towards the ends; on a flat world it would look less curved.


Horizon
Pole Star
N
S
Equator
a)
Horizon
Pole Star
N
S
Equator
b)
Pole Star
23.5°
Horizon
Pole Star
N
Tropic of Cancer
S
Equator
c)
Pole Star
N
Horizon
S
Equator
d)

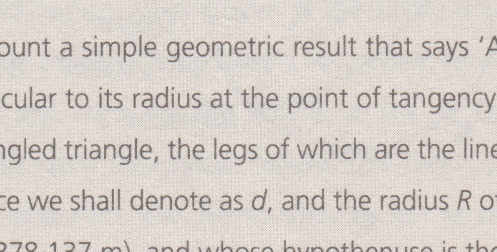
Someone travelling in the southern hemisphere will not be able to see the Pole Star, also known as Polaris (a). However, if they move northwards there will come a moment when the star appears on the horizon. This will happen just as the traveller reaches the Equator (b). If they continue travelling northwards they will see how the Pole Star gradually rises in the sky; thus, on the Tropic of Cancer, the parallel of latitude 23.5°, the Pole Star is at 23.5° over the horizon (c). And at the North Pole, the traveller would have Polaris right over their head (d).

But without doubt the most tangible and indisputable evidence for our purposes is what follows. Let's suppose that from the beach we saw a sailing boat sailing away. If the sea were a great plane of water we would not see it disappear instead we would see the yacht getting smaller and smaller until it was imperceptible. However, that is not what happens; in reality what we see in the distance on the horizon is first the boat's hull disappear, then the deck, the sails, and finally the top of the mast with its little flag. The reason for this effect is the terrestrial globe's curve. It is similar to when we see someone going into the distance over a hill, first their feet disappear, then the legs, the waist, the trunk and finally their head.

Furthermore, this effect we are talking about is the reason that the horizon is a distinct line between the sea and the sky. What would happen otherwise is that

HOW FAR AWAY IS THE HORIZON?

When do we no longer see the sails on the upper part of the mast of a yacht sailing out to the sea? The tool required for answering these questions is the famous, simple and ever-useful Pythagoras theorem: 'Given a right-angled triangle with legs (or catheti) a and b and hypotenuse c, then $c^2=a^2+b^2$.'



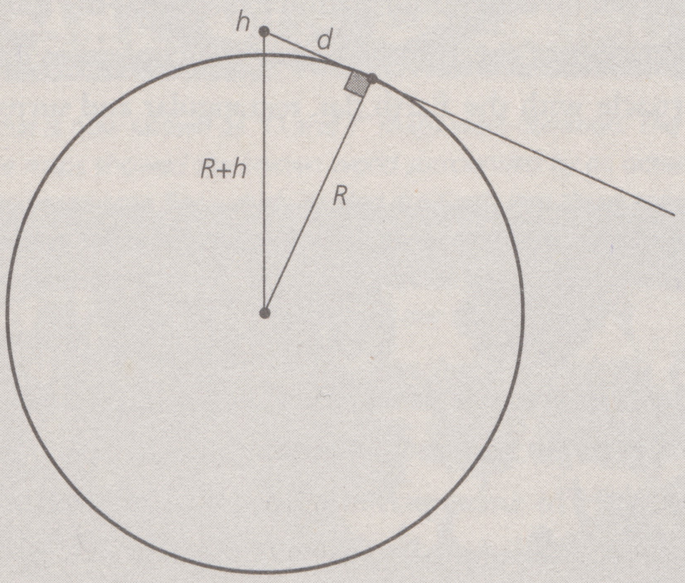
Let's first look at the question of the distance to the horizon, for which we shall suppose that the person looking at the line separating the sky from the sea has their eyes at height $h = 1.70$ m. As light travels in a straight line, when we are looking towards the horizon the line of sight is a straight line tangent to the Earth – taking into account a simple geometric result that says 'A straight line tangent to a circumference is perpendicular to its radius at the point of tangency' (see the diagram opposite). We then have a right-angled triangle, the legs of which are the line of sight from the eyes to the horizon, with a distance we shall denote as d, and the radius R of the Earth (we'll consider the equatorial radius as 6,378,137 m), and whose hypothenuse is the line that goes from the eyes to the centre of the planet, $R+h$. By using the Pythagoras theorem we get the result that the distance to the horizon is almost 5 kilometres:

$$d=\sqrt{(R+h)^2-R^2}=4{,}656.79\text{m}.$$

the area between the sea and the sky would be hazy and we would not be able to distinguish the horizon.

Mediaeval thinking

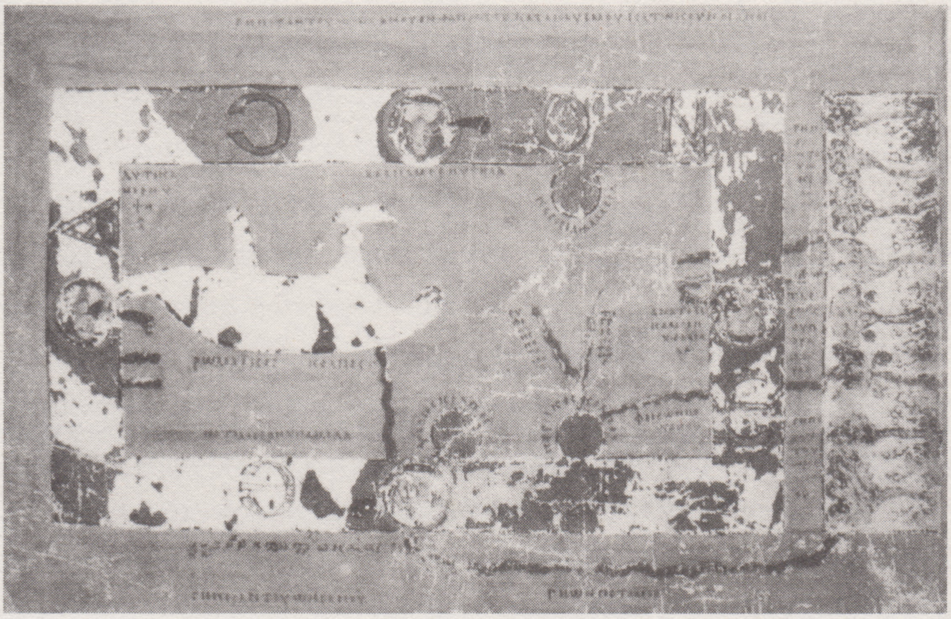
In spite of all the above, one of the beliefs that has become popular in the Western world is that in the Middle Ages everyone thought the world they lived in was flat and that it was Christopher Columbus (1451-1506) who convinced his contemporaries that it was round. This myth seems to have its origin in the book *The Life and Voyages of Christopher Columbus* by the American writer Washington Irving (1783–1859).



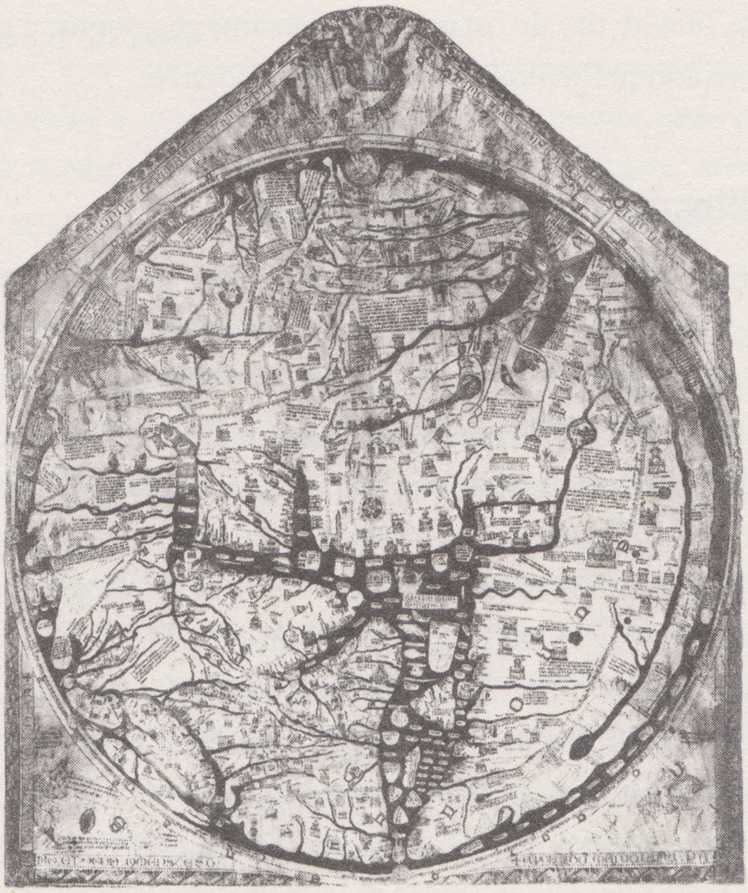
A rectangular triangle whose legs are the line of sight from the eyes to the horizon (distance d*) and the radius* R *of the Earth, and whose hypotenuse is the straight line that goes from the eyes to the centre of the planet,* R + h.

If we use the same reasoning for the top sail on the boat's mast, which we shall take to be $h = 15$ m tall, the result obtained is that a sailor in the crow's nest will see the horizon at a distance of 13,832.73 m. By adding together both results, at the exact moment that we cease to see the mast, the boat will be at a distance of 18,489.52 metres.

A literal interpretation of the Bible supposedly led to the conception of a flat Earth. For example, Daniel 4:8 says "The tree grew and became strong until its crown reached the sky, and it could be seen from anywhere on Earth." and Daniel 1:35, "The stone became a great mountain and filled the whole Earth." This would not have been possible if the Earth were not flat. Samuel 2:8 and Job 9:6 speak of the pillars supporting the Earth. But what's more, the literal interpretation of the Bible also determined the shape of mediaeval maps. They were rectangular in keeping with the words of Isaiah 11:12 or of Revelation 20:8, "The four corners of the Earth", or circular or even oval in accordance with the phrase "the circle of the Earth" in Isaiah 40:22, and normally with Jerusalem in the middle, as suggested in Ezekiel 5:5. These ideas, as well as the establishment of a general cosmological system that would replace the ideas of Ptolemy and his predecessors, were popularised by the famous *Topographia Christiana* by the Greek monk Cosmas Indicopleustes (6th century). The flat shape of the Orbis Terrarum became an official truth defended by many Christian theologians and persons close to power, and was the belief of many of the humbler people who received no other information except that the Earth was flat. Following a literal interpretation of the Bible, Cosmas Indicopleustes describes the Universe as a great tabernacle with the Earth flat, rectangular and surrounded by the ocean.



World map by Cosmas Indicopleustes. Oriented with the north at the top, the map shows the Earth as rectangular, placed in the centre and surrounded by the ocean. On the left we can see the Mediterranean Sea, with the Nile flowing into it, which is born in the ocean. On the right we can see, at the top, the Caspian Sea, and below, the Arabian Sea, Red Sea and Persian Gulf. The rivers Tigris and Euphrates flow into that latter gulf.



In the Middle Ages what is now known as "O and T" maps were common. The name comes from Orbis Terrarum. *These maps showed the known world surrounded by an ocean shaped as an O, while the T was the Mediterranean Sea, which divided the Earth into three zones. Above was Asia; to the left, Europe, and to the right, Africa. Some were very simple, while others became more complex, such as the Hereford map (above), by Richard of Haldingham, or the Ebstorf map by Gervase of Tilbury, both from the 13th century.*

But the belief in the spherical shape of the Earth was not abandoned by educated people of the times. And so we can find references to it, for example, in the work of one of the fathers of the Latin church, Saint Augustine of Hippo (354–430), who never doubted that the Earth was round, although he did doubt the presence of inhabitants in the antipodes. It is mentioned also in the encyclopaedic work *Etymology*, by Saint Isidore of Seville (c. 560–636), the Archbishop of Seville, which gathers together all the knowledge of the epoch and was one of the books most widely used in educational institutions. However, there is particular emphasis on it in Isodore's *De astronomia seu naturam rerum*, and in *The Divine Comedy*, by the Italian poet Dante Alighieri (1265–1321), and in the book *De sphaera mundi*, by the English astronomer John Halifax of Holywood (1195–1256), probably the most important textbook of all time and in print for five centuries. Halifax of Holywood gathered the knowledge set out in Ptolemy's *Almagest*, got rid of the technical terms and added new knowledge, all with the aim of making Ptolemy's geographical and cosmological knowledge more comprehensible for his contemporaries.

From the ellipsoid model to the geoid

The perfect sphere remained the model for describing the shape of the Earth up to the 17th century. As a consequence of his law of universal gravitation, the English physicist and mathematician Isaac Newton (1643-1727) came to the conclusion that it must be slightly flattened at the poles and broadened at the Equator. The centrifugal force generated by its rotation is greater at the Equator and gradually decreases until it is null at the poles; therefore it counters gravity, resulting in gravity being lesser on the central parallel. Consequently, a more appropriate model for describing our planet would be that it is an ellipsoid of revolution.

But Newton's theory that the Earth was slightly flattened at the poles was not shared by all the scientists of the time. For instance, the measurements carried out by the Italian mathematician and astronomer Jean-Dominique Cassini (1625-1712), director of the Paris Observatory, and by his son Jacques Cassini (1677-1756), on different parts of the meridian itself, led them to believe that the Earth was, instead, lengthened at the poles and squashed at the Equator. This discrepancy set off a heated polemic which ended up pitting English science against French science, and divided the Paris Académie des Sciences into two irreconcilable factions. To put an end to the controversy, around 1735 the Academy proposed sending two expeditions to measure the arc corresponding to the degree of latitude close to the pole and close to the Equator. Maupertuis (1698-1759) and Clairaut (1713-1765) went to Lapland, and Godin (1704-1760), La Condamine (1701-1774) and Bouguer (1698-1758), with the help of Spaniards Jorge Juan (1713-1773) and Antonio de Ulloa (1716-1795), went to what was then known as Peru. The results of the measurements were that Newton was right. Voltaire, a supporter of Newton, would say, referring to Maupertuis, "He has flattened the Earth and the Cassinis."

Since then, and right up to modern times, different measurements have been taken to obtain the ellipsoids that best approximate the shape of the terrestrial surface. The latest ones to be carried out, using satellite technology, were the GRS (Geodetic Reference System) ellipsoid in 1980, adopted by the International Union of Geodesy and Geophysics, and the WGS (World Geodetic System), in 1984, which is becoming a worldwide standard. The GPS (Global Positioning System) system uses this standard as its reference surface to compute latitude, longitude and elevation.

THE EFFECTS OF GRAVITY

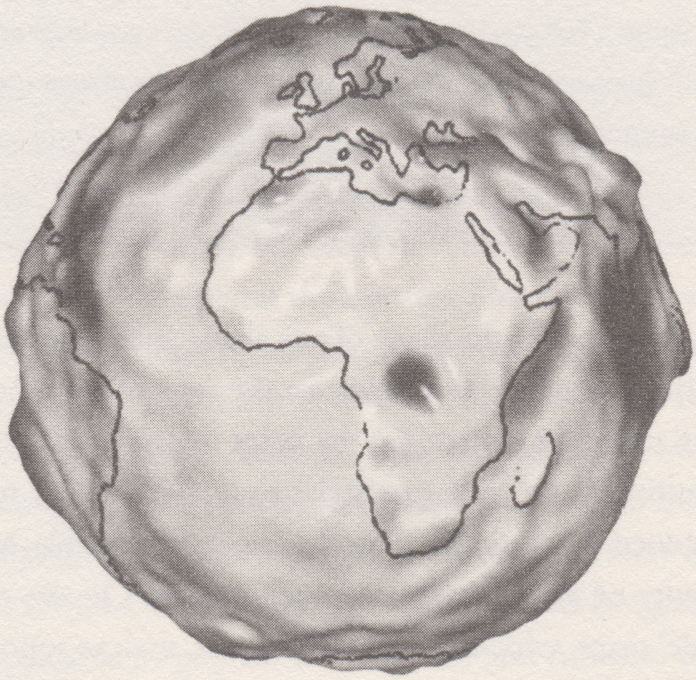
Space rockets are always launched from latitudes close to the Equator. NASA launches from Cape Canaveral in Florida; the European Space Agency (ESA) from Kourou, in French Guiana, while Russia and Japan, as they have no territory near the Equator, launch from further north and look for other solutions such as hiring launching platforms from other countries or setting up floating platforms in the Pacific Ocean. The reason for this is that there is less gravity in the zones close to the equatorial line due to the fact that the radius is greater – gravity is inversely proportional to the square of the distance to the centre of the Earth – and that the centrifugal force of the rotation of the Earth is also greater, so space craft need less fuel to get into orbit.

Gravity also has an influence in sport. In athletics such as shot-putting or the high jump, the lesser gravity in areas close to the Equator means athletes get better results than in more distant places, such as northern Europe. However, in alpine skiing, in which speed is important, the gravity in northern Europe improves the times by increasing the downhill speed compared to areas nearer the Equator.

As their approximation for the shape of our planet, cartographers use the spherical or one of the ellipsoid models, depending on the type of map they are drawing up. The sphere is the reference surface for the maps referred to as 'small scale', that is, for countries, continents or wide terrestrial areas, because for maps of such large regions the differences between both models are inappreciable, but in the ellipsoid case the complexity of the equations is much greater. On the other hand, in the case of 'large-scale' maps, which cover smaller areas, such as topographical maps or nautical charts, the differences are significant. On such detailed maps the errors in distances, areas and directions would be considerable if the sphere were taken as a reference, and so cartographers use the ellipsoid model.

It seems obvious that the shape of the Earth cannot be an ellipsoid, as it has mountains, valleys, oceans and other geographical features. In reality, what we mean when we say that the terrestrial surface is an ellipsoid is that the imaginary surface representing the average sea level spread over all points on the globe, including the emerged areas, if there were a hypothetical channel joining those points with the sea, is an ellipsoid. However, measurements carried out by geodesics show that the said surface is not an ellipsoid, but that there are variations in the level from some areas to others due to localised variations in gravity, and which are a consequence of the non-uniform geological composition and other irregularities. The geoid, a term from the Greek *geo,* 'earth' and *oide*, 'shape' (shape of the Earth), is the model that was introduced so as to take such variations into account. The geoid is the three-dimensional shape that approximates the average level of the sea, and we can think of it as the surface of the sea in calm conditions and in which at every point the direction of gravity (or plumb line) is perpendicular to the surface. In more technical terms, it is a surface with constant gravitational potential, and is used in altimetry as a reference to determine land elevation.

Throughout this book we'll consider the Earth as having the shape of a sphere; in other words, we'll take the spherical model of our planet.



Mathematical model for describing the shape of the terrestrial surface.

Chapter 2

The Dimensions of the Earth

> *"What a fool you are! Naturally I don't need to see you, if that's what you mean. You know you're not exactly a sight for sore eyes. I need you to exist and not to change. You're like that metre of platinum wire they keep in Paris or somewhere in the outskirts. I don't think anyone's ever needed to see it."*
> *"That's where you're mistaken."*
> *"Not I. Anyhow, it doesn't matter. I'm glad to know that it exists, that it measures the exact ten millionth part of a quarter of a meridian. I think about it every time they start taking measurements in an apartment or when people sell me cloth by the metre."*
>
> Jean-Paul Sartre, *Nausea* (1946)

Together with the problem of determining the shape of the planet we live on and travel around is the question of finding out just how big it is. Therefore, once it's been established that the Earth is spherical, the objective is to determine its perimeter, or its radius, as the formula of the length of a circumference (for the case of a sphere we mean here any of its maximum circles) is $2\pi r$.

The estimates of Eudoxus and Archimedes

Again it was the Greeks who provided the answer to the question. As was pointed out in the previous chapter, Aristotle, in his treatise *On the Heavens*, mentions that mathematicians had worked out the length of the terrestrial circumference at 400,000 stadia. It seems that Aristotle was quoting the work of the Greek mathematician and astronomer Eudoxus of Cnidus (c. 400–c. 347 BC), considered to be the founder of mathematical astronomy.

The next estimate of the size of our planet is to be found in *The Sand Reckoner*, a treatise written by one of the greatest Greek mathematicians, Archimedes of Syracuse

(c. 287–c. 212 BC). In the treatise he estimates the number of grains of sand that will fit into the Universe, for which he needs to calculate its size. In his calculations, he states that "the Earth's perimeter is 3,000,000 stadia and no more", although he acknowledges that in his time there were those who had estimated it at 300,000 stadia but that he believed that to be too small. Like Plato, he considered the Earth to be a very big planet.

Eratosthenes' measurement

The most famous measurement of the perimeter of the Earth in ancient times is due to Eratosthenes of Cyrene (285–195 BC), the head of the Library of Alexandria for 45 years until his death. For the task he carried out a dual measurement of the angle and of the length of an arc of the meridian that passes through Alexandria. His estimate of the complete meridian was 252,000 stadia, which, as we shall see, is a good approximation. Knowledge about his methods was passed down to us through the Greek astronomer Cleomedes (c. 10–c. 70) and some other classical authors (Hero, Strabo, Pliny, Vitruvius…). Important works from the times of ancient Greece were unfortunately lost for ever and much of the knowledge that we have on them has come down to us indirectly through other works from classical authors who commented on that era.

Eratosthenes' method took into account that the Earth is spherical, that the Sun's rays can be considered to be parallel when they reach the terrestrial surface due to the enormous distance that separates it from us, and that the cities that it

ERATOSTHENES OF CYRENE (276–194 BC)

Eratosthenes was a multitalented scholar who worked in geography, mathematics, astronomy, chronology, grammar, literary criticism and even poetry, which caused his colleagues to nickname him 'pentathlete', in reference to the athletics discipline that covers five sports. But the fact that he worked in so many fields was also the reason they gave him the epithet 'beta', in other words 'the second', which can be interpreted as a person who spends their time on many different things and therefore cannot excel in all of them. However, he was one of the greatest scholars of antiquity. At the age of 30 he was named chief librarian of the Library of Alexandria, a position he held till his death 45 years later.

deals with, Alexandria and Syene (now Aswan), were situated on the same meridian, which thus determined the meridian arc that was being measured.

The chief librarian of the Great Library of Alexandria took the distance between Alexandria and Syene as being 5,000 stadia, and thereby established the length of the arc of the meridian. To get that distance, according to the legend, Eratosthenes enquired of the travelling merchants who traded between the two cities and was told that a camel, which covered more or less 100 stadia per day, took 50 days to get from one to the other. However, it's more likely that he would have consulted the data in the library, which, being the good scientist he was, he would have compared with the information that the traders or soldiers of the Egyptian army could give him.

At the same time, Eratosthenes took into account that Syene was on the Tropic of Cancer. In other words, at midday at the summer solstice (around 21 June) the Sun's rays fall vertically on the city, which any of Syene's residents or visitors could see by noting that sunlight shines onto the bottom of a deep well.

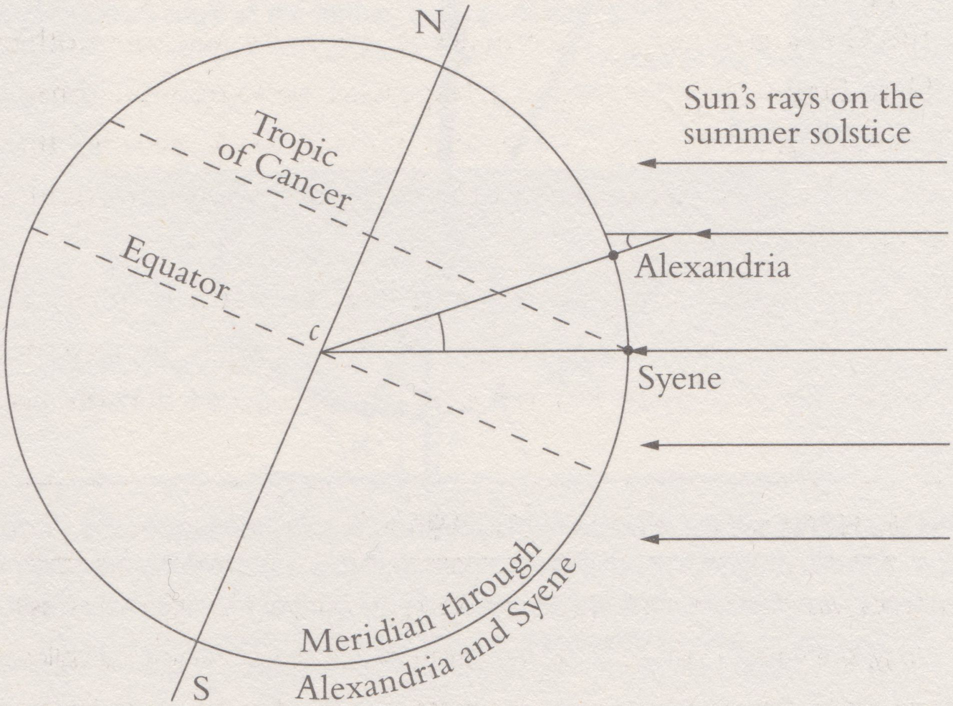
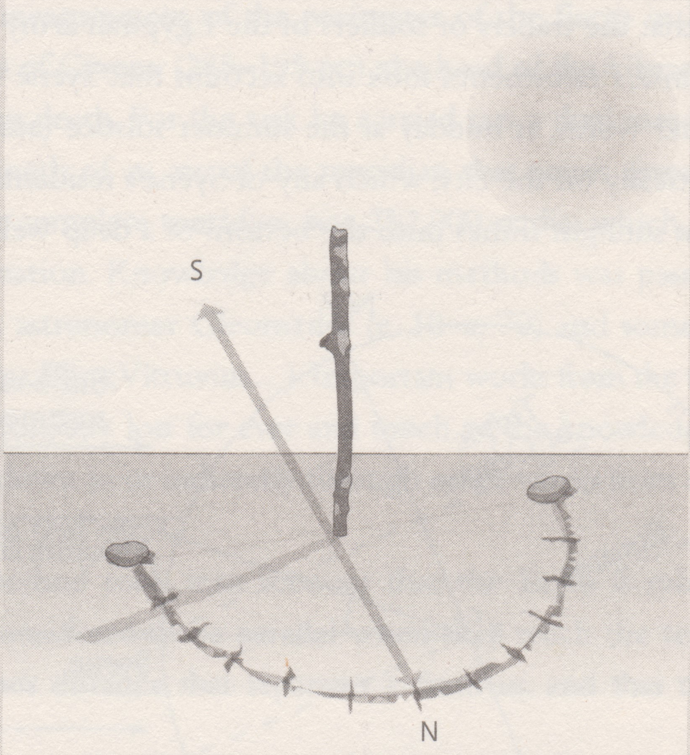

N
Tropic of Cancer
Equator
c
Sun's rays on the summer solstice
Alexandria
Syene
Meridian through Alexandria and Syene
S

Diagram of the location of Alexandria and Syene and the incidence of solar rays during the summer solstice for Eratosthenes of Cyrene's measurements of the circumference of the Earth.

In addition to the previously mentioned knowledge, to measure the angle made by the meridian arc Eratosthenes also used a gnomon, a simple device formed by a vertical stake placed on a horizontal base, which he set up in the city of Alexandria. Instead of using a simple gnomon, it is said that Eratosthenes in fact used a great obelisk.

USES OF THE GNOMON

By fixing the gnomon in a permanent position we can discover the variation of the shadow the Sun produces as it moves across the sky throughout the day. Thus it is possible to determine the moment in the day at which the Sun reaches its highest point – noon – which is the moment at which the shadow cast by the gnomon is shortest. The gnomon can also be used as a simple compass, as its shadow at midday marks the direction north–south. Specifically, for Europeans the shadow points northwards.

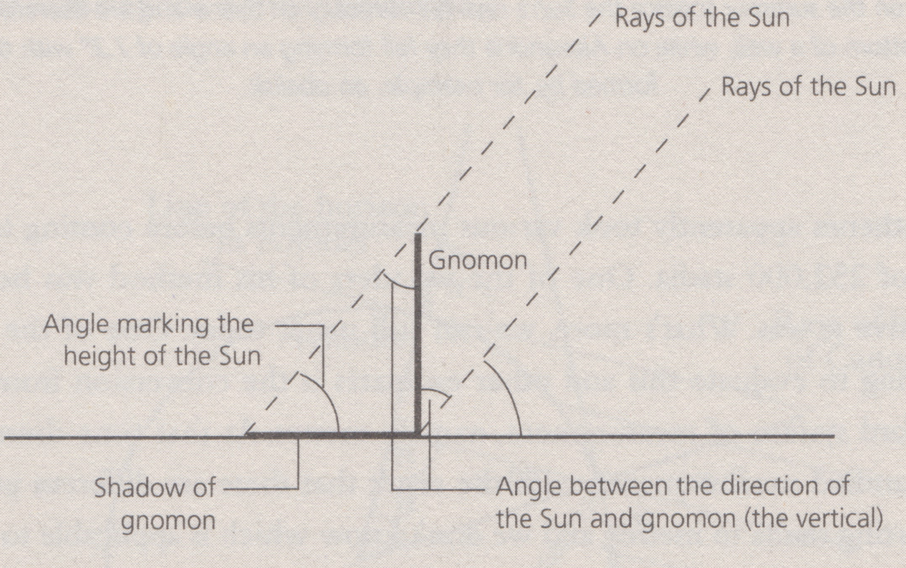


The gnomon marks the direction north–south at the solar noon, when the shadow is shortest. Throughout the day, the gnomon describes a symmetrical hyperbola in respect to the direction north–south, except on the 20th of March and the 22nd of September, when it describes a straight line in the direction east–west.

With the gnomon, Eratosthenes measured the angle α of inclination of the Sun with respect to the vertical at midday on the summer solstice, which was 7.2°, which is equal to 1/50 of 360°. At that same moment in time that the Sun's rays fall vertically on the city of Syene, then the angle at the centre of the earth projected by the meridian arc between Alexandria and Syene is exactly α, that is, 7.2°.

If we keep our gnomon up throughout the year it will also tell us the summer and winter solstices. If we make a mark at the end of the Sun's noonday shadow on every day of the year, we will see that in winter, which is when the Sun is lowest in the sky, the shadows are longest. The winter solstice is the day when the shadow is longest of all, while the shadows in summer are shorter with their apex on the summer solstice.

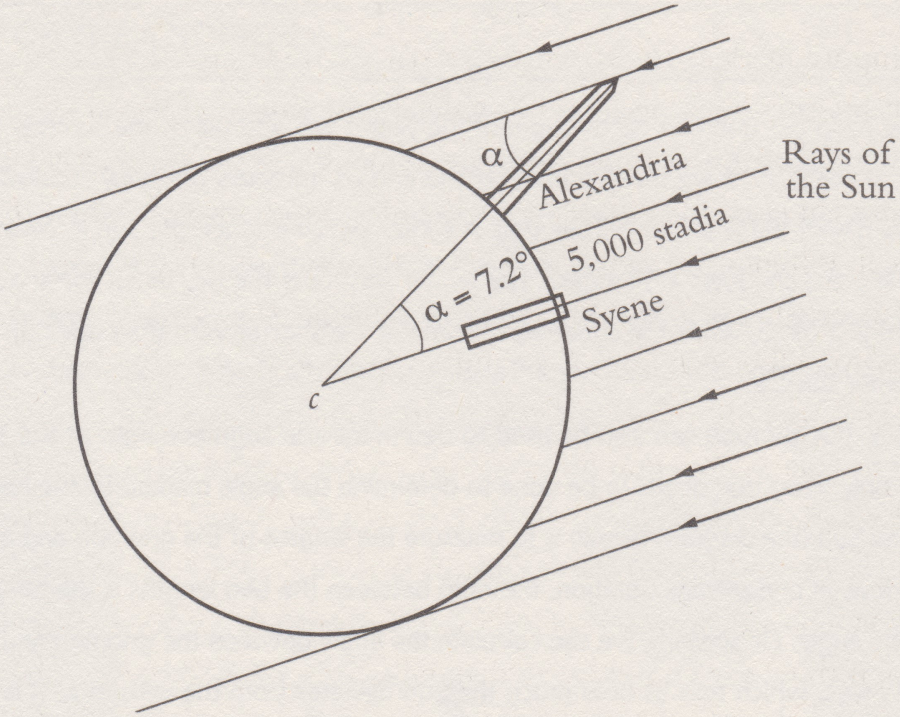
Finally, the gnomon can also be used to determine the angular height of the Sun. The only thing that needs to be done to determine the angle marked by the height of the Sun (the diagram below) is to measure the lengths of the gnomon and of its shadow. In present-day notation, the ratio between the two lengths is the tangent of the angle. Or, similarly, we can calculate the angle between the gnomon and the Sun's rays, which tells us how much the Sun deviates from the vertical as it is the complementary angle of the former; the sum of both is 90°.


Rays of the Sun
Rays of the Sun
Gnomon
Angle marking the height of the Sun
Shadow of gnomon
Angle between the direction of the Sun and gnomon (the vertical)

The gnomon and its shadow enable us to calculate the angular height of the Sun.

By means of a simple argument it can be concluded that if the meridian arc has a length of 5,000 stadia for an angle of 7.2°, then the complete circumference, that is 360°, will have a length of:

$$\frac{360^\circ}{7.2^\circ} \cdot 5{,}000 = 50 \cdot 5{,}000 = 250{,}000 \text{ stadia.}$$



At midday on the summer solstice the Sun's rays fall vertically on Syene and are therefore reflected on the bottom of a well, while on Alexandria they fall forming an angle of 7.2° with the vertical formed by, for example, an obelisk.

Eratosthenes apparently took various measurements before coming to his final estimate of 252,000 stadia. One of the wonders of his method was how simple and effective it was. What's more, we can still use it today. One of the problems when trying to evaluate this and other estimates is the conversion from stadia to our standard system of measurement, namely metres. At that time there were no standard unified measurements, with the result that there are different possibilities for converting stadia to metres, and we don't know which is applicable to the stadia Eratosthenes used. If we take the Egyptian stadium, which measured 157.5 m, his estimate would be 39,690 km, which is very close to the 40,030.2 km of the Earth's perimeter of the spherical model (obtained from ellipsoid WGS 84).

Although almost all the estimates made by the chief librarian at Alexandria were slightly incorrect, the errors compensated for each another and gave a result that was very close to the true one. For example, Syene is not exactly on the same meridian as Alexandria, and the distance between the two cities was impossible to determine at that time with any accuracy, and, with the gnomon method, the calculation of the angle between the Sun's rays and the vertical could only be worked out approximately.

Posidonius' measurement and Colombus' error

Another important estimate of the terrestrial circumference in ancient times was carried out by the Greek Stoic philosopher Posidonius (c. 130–50 BC), one of the greatest geographers of his time. His measurement has also been passed down to us through different classical authors. Again the idea was to measure a meridian arc, this time between Rhodes and Alexandria. From his observatory in Rhodes, Posidonius noted that a certain star (Canopus, the second brightest star in the sky) was situated on the horizon. However, from Alexandria the same star was seen up in the sky at an angle equal to 1/48 of the terrestrial circumference (see the next illustration). According to Cleomedes, Posidonius considered that the arc between Rhodes and Alexandria was 5,000 stadia long, which gave an estimate of the circumference of the Earth of $48 \cdot 5{,}000 = 240{,}000$ stadia. However, the Greek geographer and historian Strabo (63 BC–24 AD) quotes Posidonius's later estimate of 180,000 stadia (that is, 28,350 km, if they were Egyptian stadia), which he worked out by carrying out a new and more accurate calculation of the distance between the two cities of 3,750 stadia, thus reducing the Earth's dimensions.

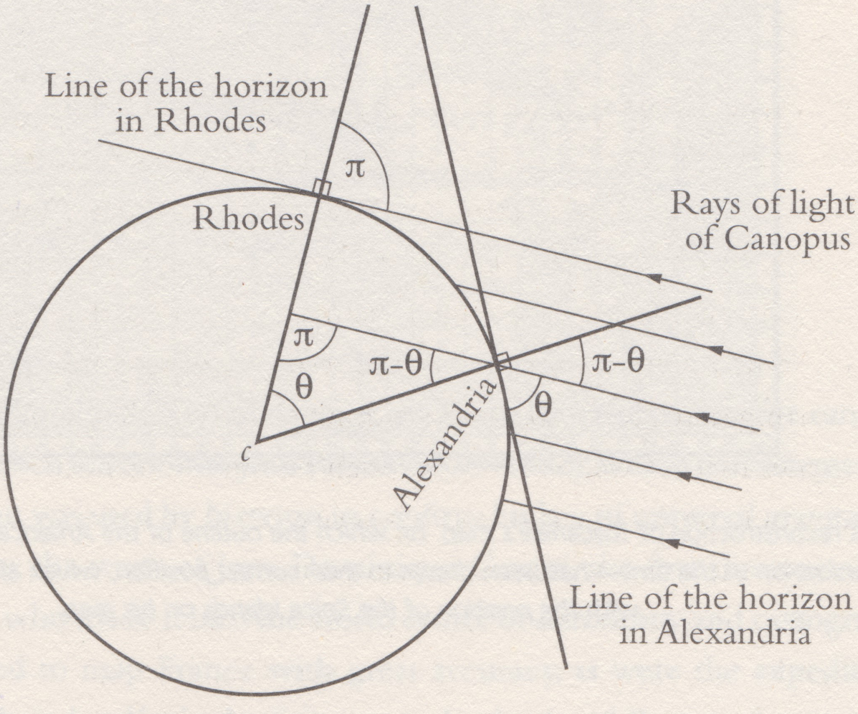
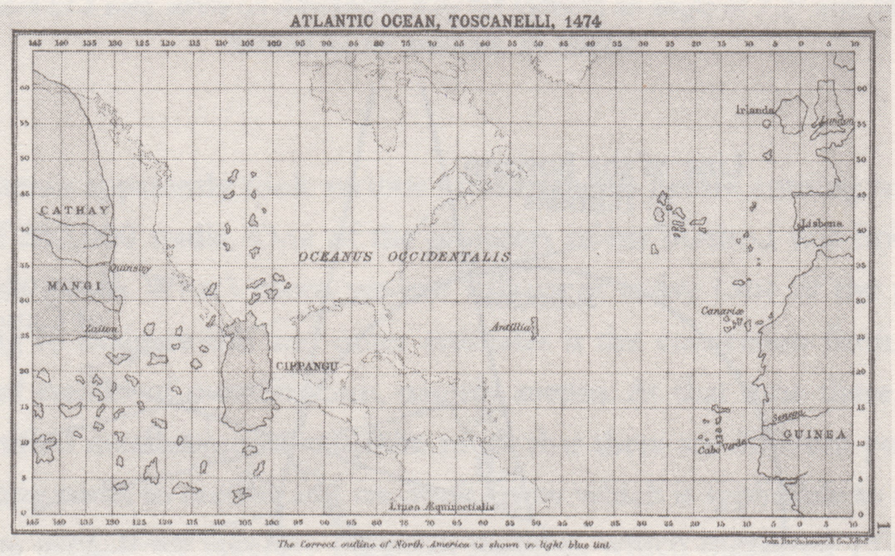

Line of the horizon in Rhodes
π
Rhodes
Rays of light of Canopus
π
π-θ
π-θ
θ
θ
c
Alexandria
Line of the horizon in Alexandria

Diagram of the measurement of the Earth's perimeter carried out by Posidonius. When the star Canopus can be seen on the horizon from Rhodes, at Alexandria it is seen at an angle of θ, equal to the angle between Rhodes and Alexandria.

The Posidonius method for calculating the terrestrial circumference is also simple and ingenious, and geometrically impeccable, though it does make an error from the point of view of physics: Posidonius did not take into account the refraction caused by the terrestrial atmosphere, which makes celestial bodies near to the horizon appear higher than their true position. If light rays were not subject to this effect we would not see Canopus at that time but some time later. Consequently, the true angle was smaller than the one calculated by the Greek philosopher.

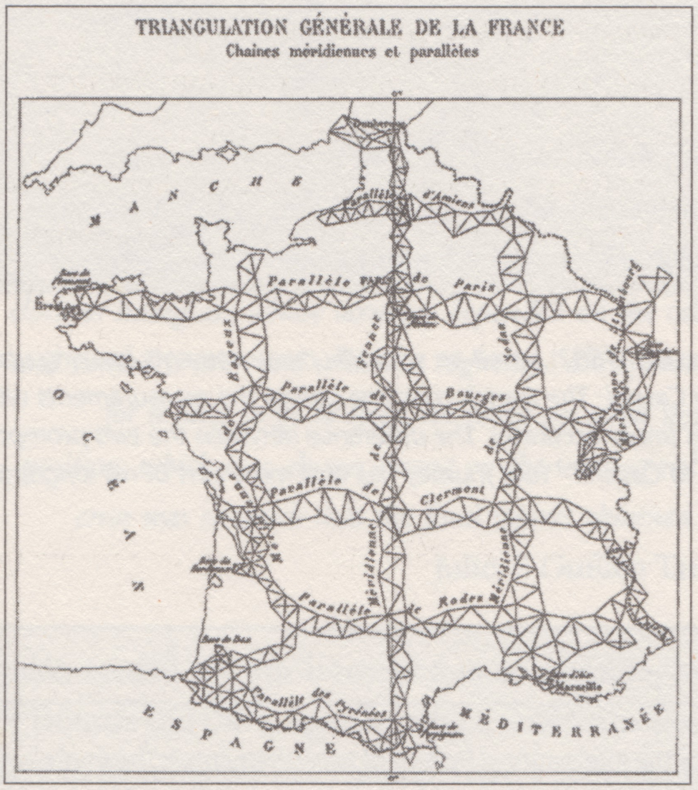
Ptolemy, like Strabo and others, accepted Posidonius's estimate as valid and included it in his *Geography*, and so popularised the idea of an Earth of small proportions, particularly among geographers and cartographers. This same perception lasted up to the 15th century, encouraging those such as the mathematician and astronomer Paolo Toscanelli (1397–1482) – who was to make a nautical chart of the Atlantic Ocean – to believe that it was possible to sail from Europe to Asia, and Christopher Columbus himself came to believe that it was feasible to open a new route for the trade in spices via the Atlantic Ocean.


ATLANTIC OCEAN, TOSCANELLI, 1474
CATHAY
MANGI
Quinsay
Zaiton
CIPPANGU
OCEANUS OCCIDENTALIS
Antilia
Irlanda
Lisbona
Canariæ
Cabo Verde
Senegal
GUINEA
Linea Æquinoctialis
The Correct outline of North America is shown in light blue tint

A reconstruction of Toscanelli's map, on which the outline of the American continents – unknown at the time – has been drawn in their correct position, which almost coincides with the position of the Spice Islands on his map.

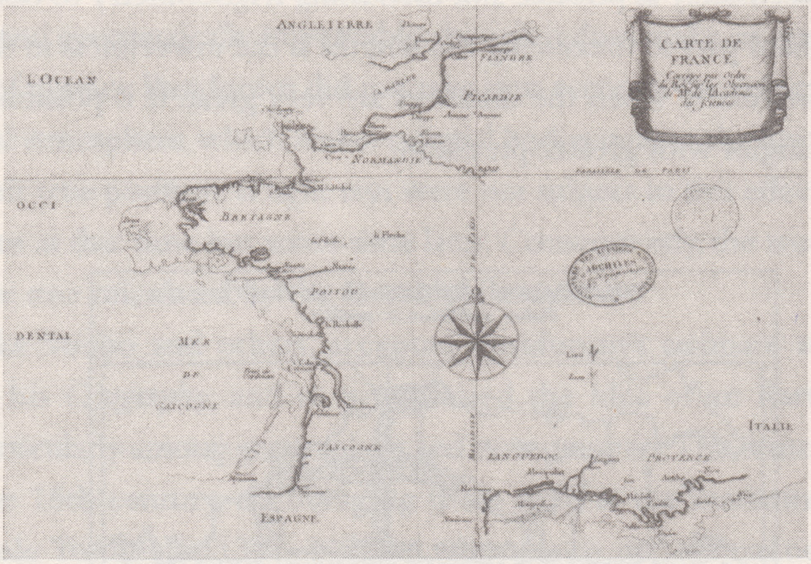
The triangulation technique

The technique that was later used for measuring the Earth's meridian, and therefore its dimensions, was triangulation. The process consisted of dividing the land into triangles, measuring the triangulation angles and the length of one of the sides, called the 'base', as accurately as possible and then using trigonometry to obtain the other lengths. The hard part in measuring the lengths is the unevenness in the land being measured, particularly when the distances are very great. It is possible, however, to measure the angles with greater precision.


TRIANGULATION GÉNÉRALE DE LA FRANCE
Chaines méridiennes et parallèles
MANCHE
Parallèle d'Amiens
Parallèle de Paris
Parallèle de Bourges
Parallèle de Clermont
Parallèle de Rodez
Parallèle des Pyrénées
Méridienne de France
OCÉAN
ESPAGNE
MÉDITERRANÉE

Above, a general triangulation of France carried out between 1818 and 1845.

Among the documentation on measuring the Earth by triangulation we can find the work of the French astronomer Jean Picard (1620–1682), whose measurement of the Earth's meridian was used by Newton to confirm his law of universal gravitation. Also important was the work of Jean-Dominique Cassini, the first director of the Paris Observatory, who made it into the world centre of astronomy and cartography and later attempted to map France with great accuracy, as were the expeditions organised by the Paris Académie des Sciences to Lapland and Peru to determine if the planet was flattened or elongated at the poles. The measurement of the meridian between Dunkirk and Barcelona, carried out by the French scientists Jean-Baptiste Delambre (1749–1822) and Pierre Méchain (1744–1804), led to the determination of the size of a metre.



A map of the French coasts (1682) based on scientific measurements (making use of triangulation) by Picard, Lahire and Cassini. The French coastline before the measurements were made (larger area) and afterwards (more accurate). The difference between the two prompted Louis XIV to remark to Cassini: "Your journey has cost me a part of my kingdom!"

THE METRE

The unit of length of the International System of Measurements is the metre, which nowadays is defined as the length of the distance covered in a vacuum by light in 1/299,792,458 seconds (approximately 3.34 nanoseconds, that is, 3.34 billionths, 10^{-9}, of a second).

The definition of metre has changed over time, but its inclusion as a universal measurement is due to the French Revolution. The Commission of Weights and Measures was set up in 1790 to standardise and unify units of measurement. Two conditions were set: the measures must be universal, acceptable to everyone and not arbitrary. For that reason the new unit of measure, the metre, was defined as the ten millionth part of the quadrant of the meridian between the North Pole and the Equator. In the midst of the revolutionary unrest, a double expedition was organised to measure the meridian that passes through Paris from Dunkirk to Barcelona. Delambre, who set off for Dunkirk, and Méchain, who headed for Barcelona, experienced all kinds of adventures and vicissitudes during the long process of triangular measurement which lasted 7 years owing to the social unrest of the time. We can find a very interesting fictionalised version in *The Measure of the World*, by Denis Guedj.

Chapter 3

Meridians, Parallels and Great Circles

> *From the height of the Sun and the North Star, the latitude could be reckoned; the chart and the sailing needle, the compass, with speed reckoned by sight and the time measured – if that's the right word – with the sand clock or hourglass, whose accuracy depended on the cabin boy who watched it and turned it over, and which at nighttime always went fast, eating up sand, for it later to go to sleep, we could get what now we call the estimated position, which was always wrong on the wind speed, and in those days the error was so much that the final figures obtained could be called fantasy.*
>
> Julio F. Guillén Tato, *La Náutica* (1935)

The next element needed for our introduction to cartography is the geographical coordinates, the latitude and the longitude, which enable the position of any point on the Earth's surface to be uniquely determined. Here, a geographic grid formed by two notable families of spherical curves, the parallels and the meridians, the sphere's constant curves of latitude and longitude, plays its part. While the belief that a straight line is the shortest route between two points seems so natural to us, it can be difficult for us to take on board that on the terrestrial surface that cannot be right, the reason being that there are no straight lines on it. The next question seems evident: what are the curves on the sphere that take the role of the straight lines on the plane? And, more to the point, what is the shortest path between two points on a spherical surface? We shall see that the answer is another interesting family of spherical curves, the great circles.

Latitude and the parallels

To be able to define the geographic coordinates we have to take into account the rotational movement of the Earth on an imaginary axis which passes through its centre. The North and South Poles are the intersection points of that axis with the terrestrial surface, and the only points that remain fixed under the rotational movement. In the case of the spherical model of the planet, the parallels are the circumferences obtained by intersecting the sphere with planes perpendicular to its axis of rotation (see next image). There is a parallel that stands out, the Equator, which is halfway between the North Pole and the South Pole. This is obtained by intersecting the globe with the plane perpendicular to the axis passing through the centre of the sphere. What's more, it is the parallel of greatest length.

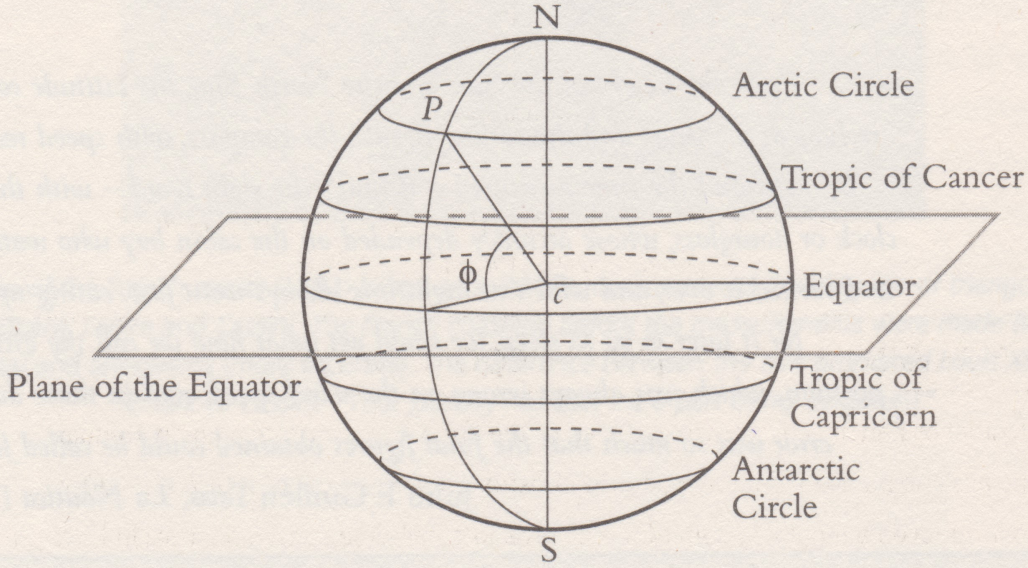
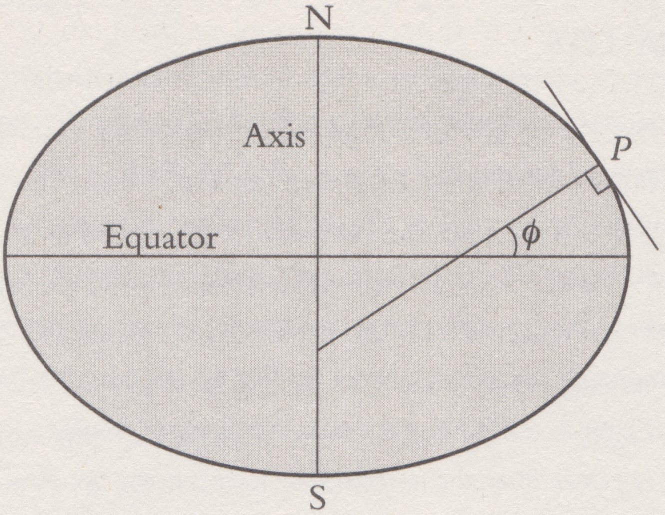


Diagram showing the five principle parallels and the latitude of a point P.

The latitude of a point on the terrestrial surface is defined as the angle of inclination in respect to the plane of the Equator, in other words the angle between the segment joining the centre of the Earth with our point and the plane on which the Equator lies (an angle denoted by ϕ in the previous image). The latitude of Bilbao is 43° 15' 52" N, that is to say, it is 43 degrees, 15 minutes and 52 seconds to the north of the Equator. Latitude takes values of between 90° S and 90° N or, expressed mathematically, between –90° and 90°. The parallels are, therefore, the curves formed by the points on the same latitude.

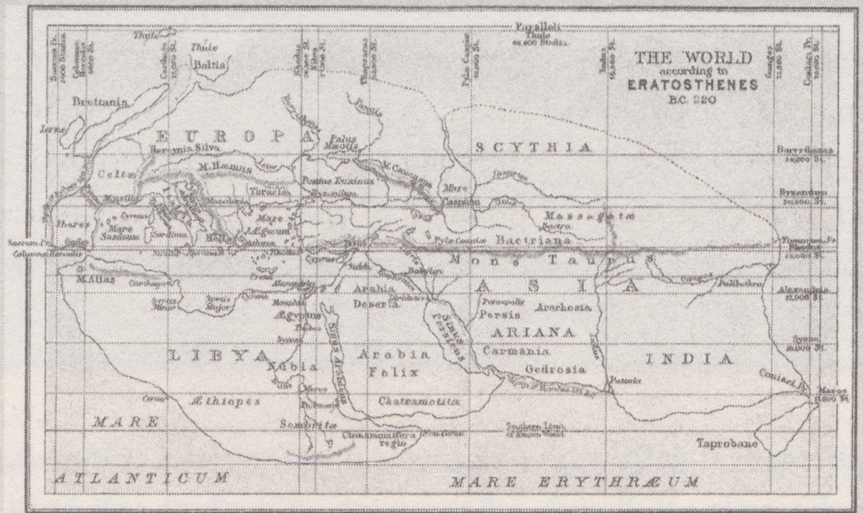
The definition of latitude we have provided is valid for the spherical model of the Earth, which is the one we are working with in this book. For the ellipsoid model, however, it is necessary to introduce a more general definition, the geodesic latitude, which is defined as the angle between the perpendicular to the straight line tangent to the meridian of the ellipsoid that passes through the point and the equatorial plane (above right).


N
Axis
P
Equator
ϕ
S

The geodesic latitude generalises the concept of latitude in the ellipsoid model of the Earth's surface.

THE ORIGIN OF GEOGRAPHICAL COORDINATES

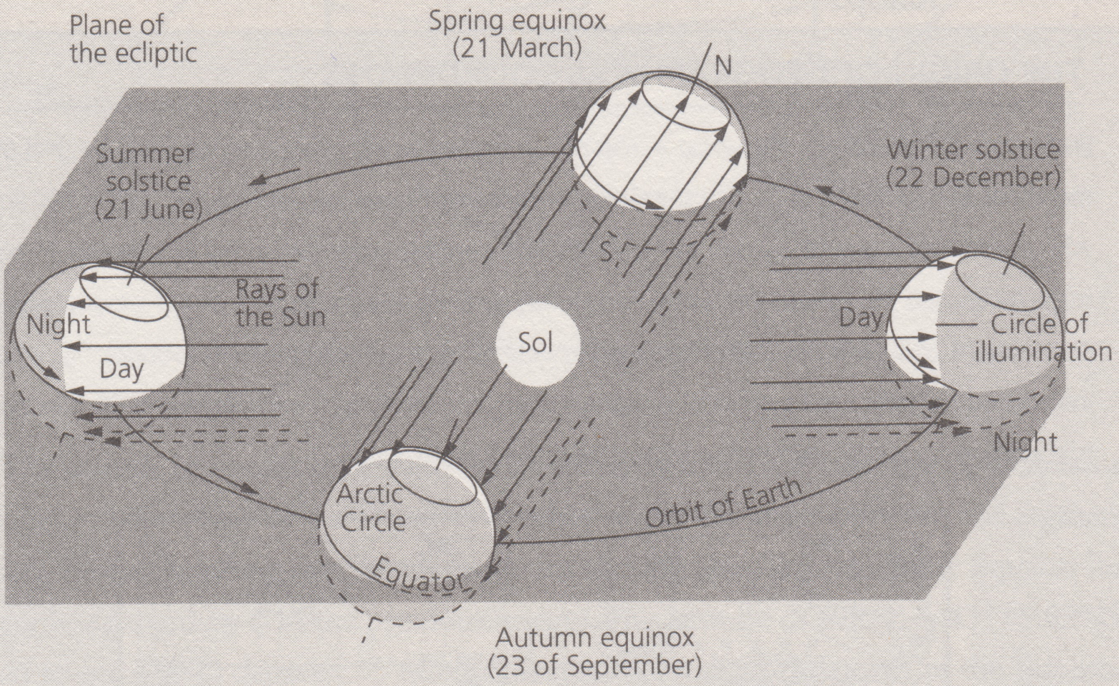
Eratosthenes had already drawn up a map of the known world containing an irregular grid of meridians and parallels, but it was the Greek astronomer Hipparchus of Nicea (c. 180–c.120 BC) who introduced a regular system of meridians and parallels separated at equal intervals. In his maps he divided the inhabitable world into eleven parallels and proposed that the simultaneous observations of lunar eclipses could be used to determine longitude. It was also Hipparchus who introduced to Greece the system of dividing a circle into 360°, with each degree in 60 minutes and each minute in 60 seconds, a legacy from the Babylonians.


THE WORLD
according to
ERATOSTHENES
B.C. 220
Paralleli Thule
Thule
Baltia
Brettania
Ierne
EUROPA
Hercynia Silva
M. Haemus
Celtae
Iberes
Mare Sardoum
Mare Aegeum
Pontus Euxinus
Palus Maeotis
SCYTHIA
Mare Caspium
Massagetae
Bactriana
Mons Taurus
ASIA
Arabia Deserta
Persis
Arachosia
ARIANA
Carmania
Gedrosia
INDIA
Arabia Felix
LIBYA
Nubia
Aegyptus
Aethiopes
Chatramotita
Taprobane
MARE ATLANTICUM
MARE ERYTHRAEUM

Eratosthenes' map, with an irregular grid of meridians and parallels.

SPECIAL PARALLELS

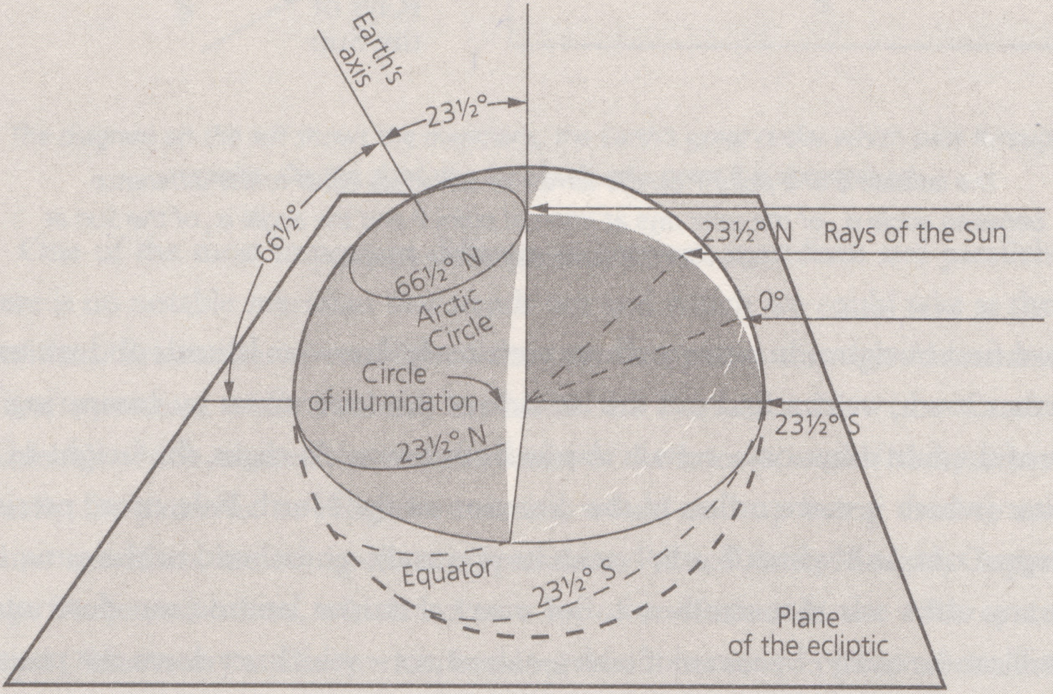
The Earth, and in particular its centre, moves around the Sun describing an elliptical orbit, very close to being a circle, which lies on a plane – called the elliptic – with respect to which the terrestrial axis is inclined at an angle of 23°30'. During this movement there comes one day in the year (around 21 June) when the Earth's axis is pointing towards the Sun, the northern hemisphere is nearer to the Sun and it has its longest day of the year (while it is the shortest in the southern hemisphere). This is the summer solstice. At its noon, the Sun is situated exactly over the parallel of latitude 23°30' N, known as the Tropic of Cancer. On the other hand, at the winter solstice (22 December) the terrestrial axis points in the opposite direction to the Sun and it is the shortest day of the year in the northern hemisphere. The Tropic of Capricorn is the


Plane of the ecliptic
Spring equinox (21 March)
N
S
Summer solstice (21 June)
Winter solstice (22 December)
Rays of the Sun
Night
Day
Sol
Day
Circle of illumination
Night
Arctic Circle
Equator
Orbit of Earth
Autumn equinox (23 of September)

A diagram of the movement of the Earth showing the solstices and equinoxes.

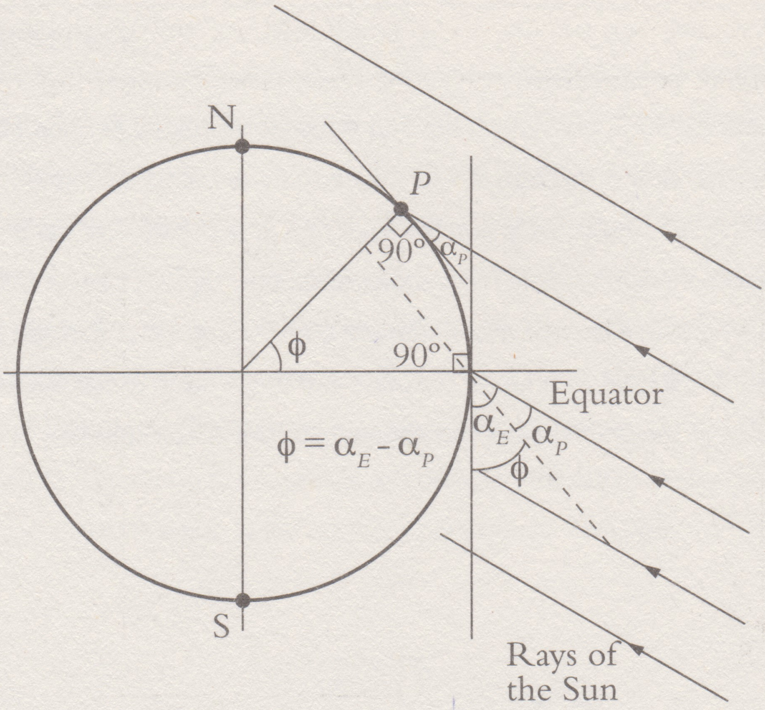
The mathematical definition of latitude is correct and comprehensible, but how can we determine it directly when we are on the high seas or in the middle of a land journey? Nowadays the solution is to use GPS, which has become a very useful tool. But let's bear in mind that this is a very recent invention, and that humanity has had to use more natural solutions throughout history.

parallel of latitude 23° 30' S, over which the Sun is precisely situated at noon on the winter solstice. At autumnal and vernal equinoxes, the Earth's axis points completely to the right or to the left of the Sun, respectively, and the solar rays at noon are over the Equator. As at the summer solstice the Sun's rays fall perpendicular on the Tropic of Cancer (23° 30' N), the places on the planet with a latitude that is at more than 90° from that parallel, to the south of parallel 66° 30' S, will have 24 hours of darkness. But to the north of parallel 66° 30' N they will have 24 hours of sunlight. The opposite happens at the winter solstice.


Earth's axis
23½°
66½°
66½° N
Arctic Circle
Circle of illumination
23½° N
23½° N
Rays of the Sun
0°
23½° S
Equator
23½° S
Plane of the ecliptic

At the winter solstice it's nighttime 24 hours a day further north than parallel 66°30' N, the Arctic Circle.

To determine terrestrial latitude it was taken into account that the angle ϕ is equal to the difference between the angle of the Sun at noon at the point at which we want to calculate the latitude and the angle of the Sun at noon on the same day on the Equator (for which we can make use of the gnomon, for example).

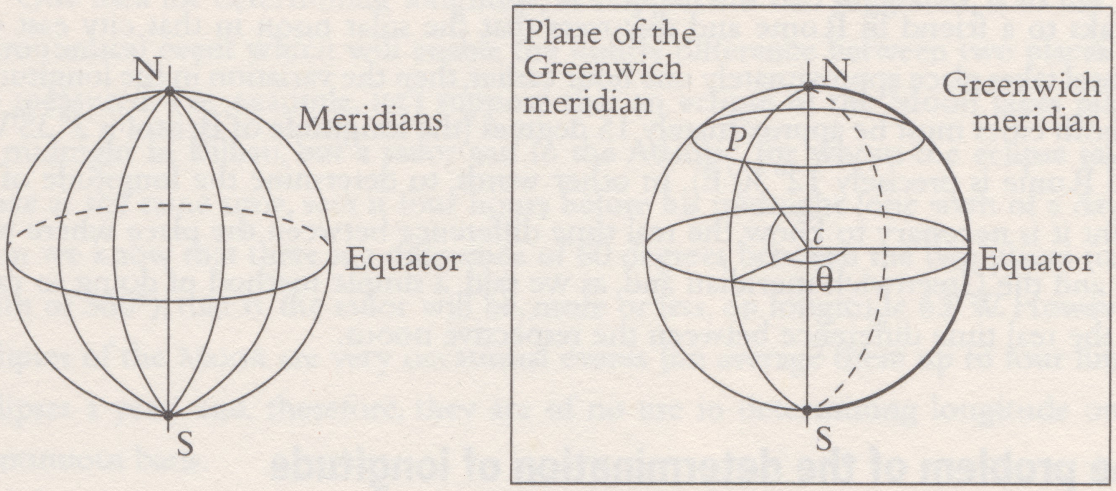


The latitude ϕ of a point P on the terrestrial surface is equal to the difference between angle α_P of the solar rays at noon at point P and the angle α_E of the Sun at noon on the same day at the Equator.

If we find ourselves in a town where we already know its latitude and we travel to another town, we can find out the latitude of this latter town by comparing the angles of the Sun's rays on a certain day in both towns. At night, the height of the Pole Star (which is more or less in the direction of the North Pole, more precisely one degree out, and which hardly varies its position) can be used in the same way, as can any other star that stands out. For many centuries latitude was determined (and still can be today) by mean of tables – almanacs – which set down the position of the Sun and other heavenly bodies with their dates and times, and by using instruments that enable the angle of height of heavenly bodies to be measured, such as the astrolabe, the quadrant and the cross staff (later to become the sextant).

Longitude and the meridians

While latitude measures our position in the direction north–south, longitude does it in the direction east–west. As a first step, we shall look at the circumferences that are obtained as intersections of the Earth's sphere with the planes that contain its axis of rotation (next image). The meridians follow each of the semicircles determined by the poles. The solar noon occurs down the length of each of them at the same time (the word *meridian* comes from the Latin term *meridies*, meaning 'midday'.



The diagram on the left shows the meridians, the Earth's great circles which pass through the poles. On the right, the longitude of a point P *is determined.*

One of the most important differences between meridians and parallels is that there is no notable meridian that stands out and which we could take as the prime meridian. Erastosthenes established the one that passes through Alexandria as his prime meridian, while Ptolemy set his in the Fortunate Isles (the Canaries and Madeira), which marked the western limits of the known world. Many other places, usually for patriotic or religious reasons, would end up being used to mark the prime meridian, including Mecca, Jerusalem, Paris, Rome, Madrid, Copenhagen, Cape Verde and others, which led to a lot of confusion. In the 18th century, the Royal Observatory at Greenwich, London, would become an international reference after the publication in 1767 of the most exhaustive nautical almanac published so far, the result of which was that in 1884 the International Meridian Conference in Washington, D.C., accepted the meridian line passing through east London as the reference meridian.

The longitude of a point on the Earth's surface is the angle of rotation relative to the Greenwich meridian, in other words, the angle between the meridian of that point, or more specifically the plane that forms it, and the plane on which the reference meridian lies (the angle denoted by θ in the previous image). The longitude at Bilbao is 2° 55'43"W, in other words, it is 2 degrees, 55 minutes and 43 seconds west of the Greenwich meridian. Longitude takes values between 180°W and 180° E or, expressed mathematically, between −180 ° and 180 °.

The Earth takes 24 hours to make one complete turn around its axis, in other words, to rotate 360 degrees, which means that this, the third planet in the Solar System, rotates through an angle of 15 degrees every hour. If an inhabitant of Bristol speaks to a friend in Rome and discovers that the solar noon in that city east of Bristol takes place approximately one hour earlier, then the variation in the longitude of both cities must be approximately 15 degrees (the longitude of Bristol is 2° 35'W and Rome is precisely 12° 30' E). In other words, to determine the longitude of a point it is necessary to know the real time difference between the place where we are and the Greenwich meridian and, as we said, a simple method of doing so can be the real time difference between the respective noons.

The problem of the determination of longitude

Just as in the case of latitude, we can set ourselves the problem of determining the longitude of any terrestrial point. One possible solution comes, again, from observing the sky. However, in this case the task is much more arduous since, as the day goes by - that is, as the Earth turns – the celestial bodies gradually disappear from the sky in the west as others appear in the east. Therefore, determining the east-west position by using the stars becomes a more complicated matter. In fact, the search for a solution to the problem of longitude lasted four centuries. The great seafaring nations, such as Britain, the Netherlands, Spain and France, offered huge prizes (reflecting

COLUMBUS'S FIRST VOYAGE

On 3 August, 1492, Christopher Columbus set sail on his voyage towards Asia across the Atlantic Ocean. He would first head for the Canary Isles with three ships, *La Pinta*, *La Niña* and *La Santa María*, which were about 22 metres long and 7.5 metres wide and which, altogether, carried 90 persons. From the Canary Islands they would set sail on the 6 September towards the west, following a more or less straight line (to make it simpler by following a constant latitude) between the parallels of latitude 26° and 30 °. Columbus had calculated that they would arrive at Cipango (Japan) in 25 or 30 days. However, on 12 October (which corresponds to 21 October in the present-day calendar) he came to the island of San Salvador (called Guanahani by the natives) and the fleet continued sailing in the area thinking that he was in the Asiatic archipelago, somewhere near Japan.

the importance of navigation for the development of those countries) and great scientists such as Galileo Galilei, Jean Dominique Cassini, Christiaan Huygens, Isaac Newton and Edmond Halley took part in that search. Possibly the biggest reward was the one offered by the British Government in 1714, which came to £20,000.

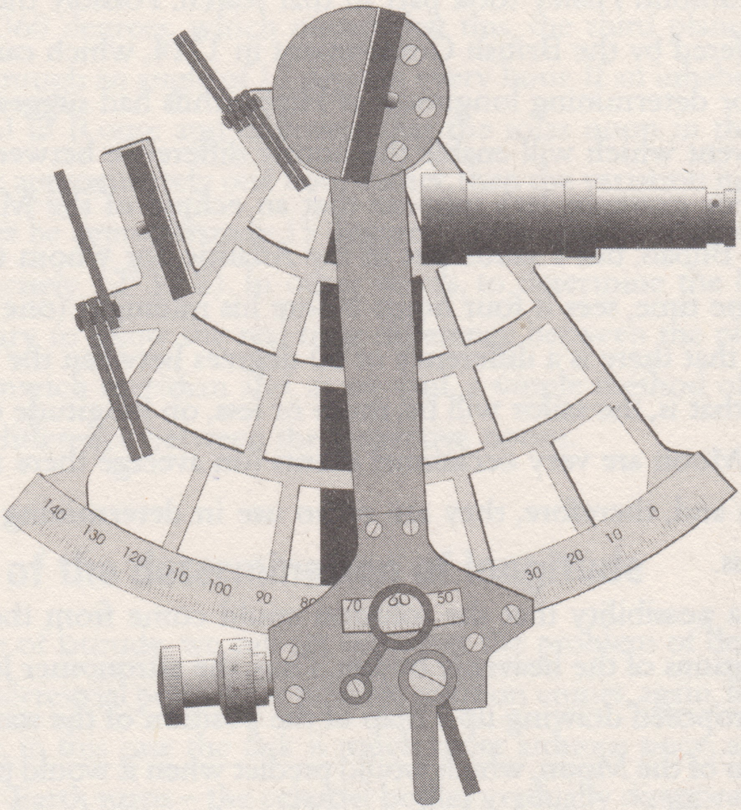
One idea for determining longitude, as Hipparchus had suggested, is to use an astronomical event which will enable the hourly difference between two places to be measured. For example, let's suppose that an eclipse of the Moon takes place at midnight in Bilbao, but a sailor out in the Atlantic, for whom the eclipse takes place at the same time, sees it four hours before his midnight (one sixth of a day) , then we know that there is a difference of 60 degrees between the two places (one sixth of 360°), that is, the sailor will be, more or less, on longitude 63°W. However, eclipses of the Moon are very occasional events (on average there up to four lunar eclipses a year) and, therefore, they are of no use in determining longitude on a continuous basis.

There was a possibility that the solution could come from the knowledge of the relative positions of the heavenly bodies. Thus, the astronomer Johannes Werner (1468–1522) proposed drawing up a map of the position of the stars in accordance with the motion of the Moon, which would predict when it would pass near to them over several years. This method, which would have been very useful for sailors, had the drawbacks that the positions of the stars were not accurately known, that there were no instruments to measure the distances between the stars and the Moon, and that the movement of the Earth's satellite was not well understood, which meant it was difficult to predict exactly where it would be.

Galileo Galilei (1564–1642) proposed taking advantage of the eclipses of Jupiter's moons, which occurred a thousand times a year in a very predictable manner, so they could be used as an astronomical clock. The idea, however, was not welcomed by the governments of the time, and there were also serious problems about getting a precise observation of Jupiter while at sea.

New methods kept being proposed, some of them outlandish and others more sensible, such as using a compass, which was based on the fact that terrestrial magnetism varies from some points to others.

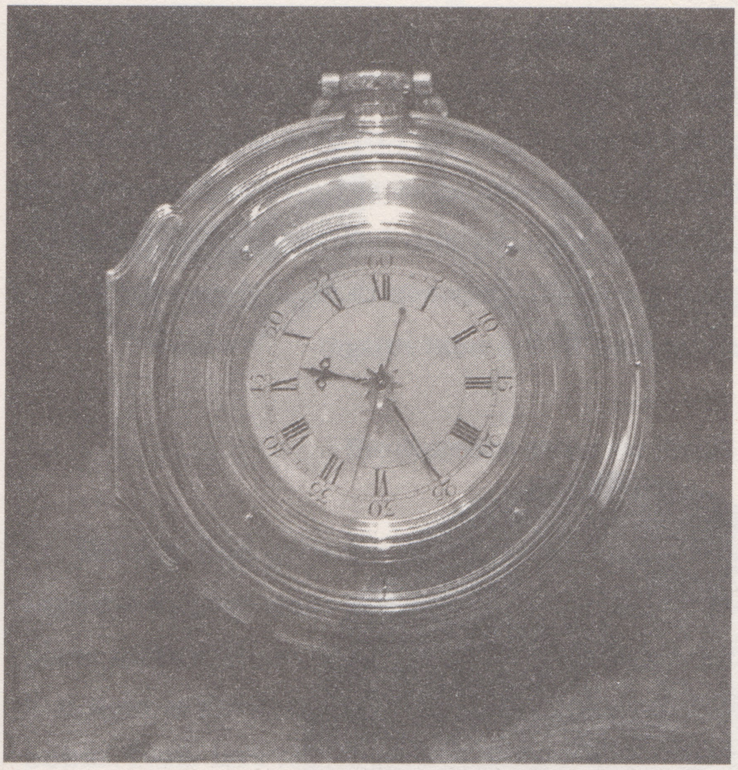
Later on, the method concerning the Moon's path and its distance from the stars would be revived. This was possible thanks to the progress made in observing the sky; to the publication of an exhaustive almanac by the Royal Observatory at Greenwich; to there now being more precise knowledge of the Moon's orbit thanks to Newton's theory of gravitation; and also to the development of better tools for sailors, such as the reflecting quadrant and the sextant.



The sextant is an important instrument in marine navigation as it allows measurements to be taken of the angles between two stars or points on the coast, or the altitude of a star.

The best solution to the problem of longitude, however, was to come from the hands of a clockmaker, the Englishman John Harrison (1693–1776), who invented a highly accurate maritime clock that carried the exact time of the port of origin to anywhere in the world, and thus enabled an exact calculation of that place's longitude. Sailors on the high seas simply needed to observe when the solar noon occurred, look at the time on the clock, which showed the time at the port of origin, calculate the variation in hours between the port and the ship, multiply the number of hours by 15 °, and they would have the variation of longitude with respect to the port of origin. This 'mechanical' solution to the longitude problem was not to the liking of scientists of the time, nor to the members of the Board of Longitude appointed by the British government, who either ignored or abused Harrison the watchmaker, and even delayed presenting the prize in the hope that a solution involving astronomy would be forthcoming. In the end, however, they had no choice but to accept, albeit very reluctantly, the evidence of the efficacy of Harrison's marine chronometers (although the only thing astronomical about them was the price).

The result of all this is that we can say that any point located on the Earth's surface is determined uniquely by the parallel and the meridian that pass through that point, or, likewise, by its latitude and longitude, which are known as 'geographical coordinates'.



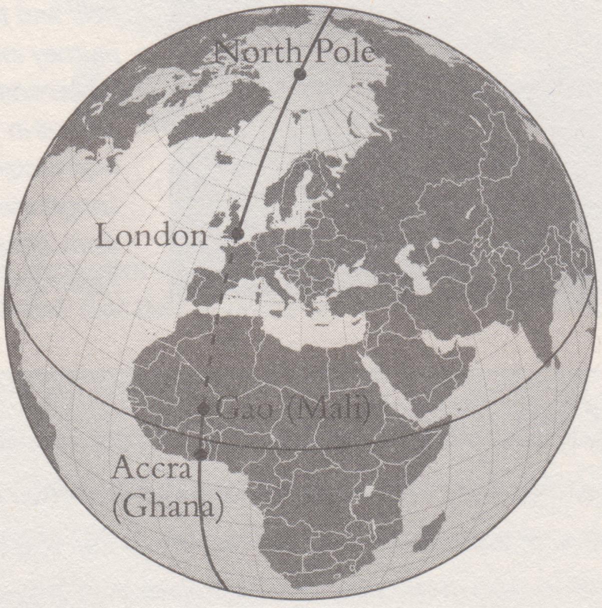
A chronometer by John Harrison, the H5. The English watchmaker's chronometer H4 solved the problem of calculating longitude. It was like a kind of large pocket watch about 13 centimetres in diameter. Proof of its efficacy was the voyage the ship named Deptford *made to Jamaica. On its arrival at Port Royal, two months later, the H4 had only lost 5 seconds. The return voyage was hell, and the total error on the journey increased to 1 minute and 54 seconds. However, the error it caused in longitude was less than that required by the Decree of Longitude. Even so, it was many years before John Harrison was rewarded with the £20,000 prize.*

THE SINKING OF THE *TITANIC*

We've all seen some of the many films and documentaries made about the sinking of the *Titanic*, and maybe through them we know that it was a ship that had been designed using the most advanced technology of the time and was unmatched in luxury. It was the pride of its owners and was to be the star of the transatlantic routes at the beginning of the 20th century. However, on the night of 14 April 1912 it collided with an iceberg and ended up at the bottom of the ocean. Knowledge of the coordinates of the location where the incident took place enabled the survivors to be rescued. The *Titanic* sent an SOS radio messages: "We have struck an iceberg. We're sinking. *Titanic*. 41° 16′ N, 50° 14′ W. Come at once." The nearest ship to pick up the message, *Carpathia*, quickly headed for the location marked by the geographical coordinates and arrived in time to rescue more than 700 people, most of them women and children in lifeboats.

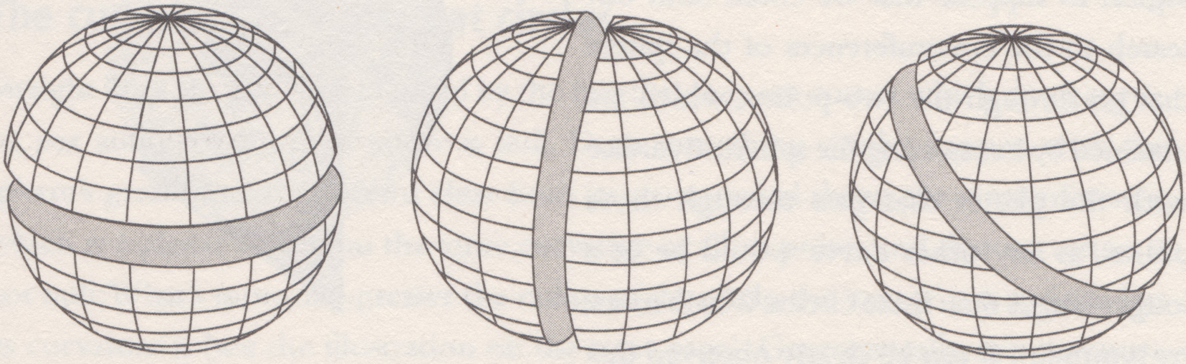
Great circles, the geodesics of the sphere

The distance between two points on any surface can be defined, and we do so in geometry, as the smallest of the lengths of all the curves (of the paths) that joins those two points, which, in short, is simply the length of the shortest route that joins them, provided one exists. In geometry we use the term 'geodesics' to refer to the curves which give us the shortest paths on a surface, although in actual fact the concept is used somewhat more widely, as it includes curves that 'locally' give us the minimum distance. What does this mean? It means that we can have two points on the surface joined by a geodesic and that the geodesic does not give us the minimum distance between them, but that when we take any two intermediate points on the geodesic, closer to each other, the shortest path between them is always the part of that geodesic joining them, as can be seen in the next image:



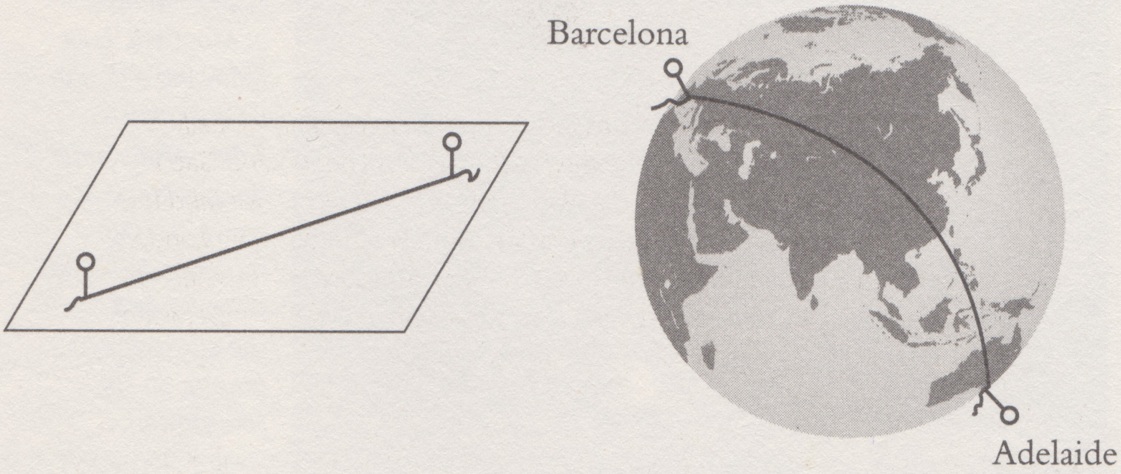
Geodesics minimise lengths locally, but not necessarily globally. For example, the part of the meridian that joins London and Gao in Mali, for instance, through the North Pole, the Atlantic Ocean and the South Pole, is a geodesic, but is not the shortest path between London and Gao. This geodesic is, however, the shortest route between nearby points as, for example, between Gao and Accra in Ghana, or between London and the North Pole.

The geodesics of a flat surface are straight lines, as we all know very well. However, on a sphere the minimum distances are obtained by means of the maximum circumferences, popularly known as great circles, i.e., the curves that are obtained as intersections of the sphere with planes that pass through its centre. Meridians are examples of great circles of the sphere. However, the only parallel that is a great circle is the Equator.



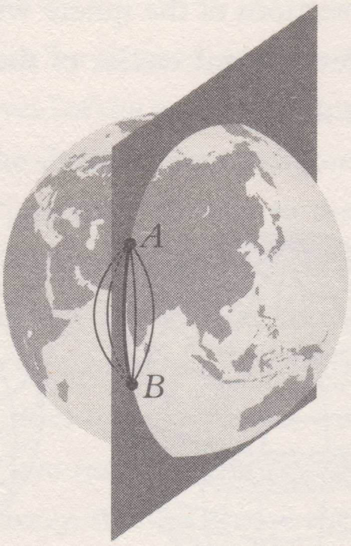
The illustration shows the Earth's great circles.

We can find primary evidence of this fact in the following experiment. Let's suppose that we want to draw the straight line that passes through two points on a plane surface. To do this we can join the points with a simple piece of string, pulling it as tight as possible so that it describes a straight line that passes through those two points. Now let's take the case of a terrestrial globe. To work out the shortest path between two points on the globe, for instance, Barcelona and Adelaide, we take a string and use it to join the cities and pull it taut. In this way we will obtain the curve of least length between the two cities (the geodesic), which is simply a great circle that passes through them, as is shown in the illustration:



The taut string seeks the shortest route between the fixed ends.

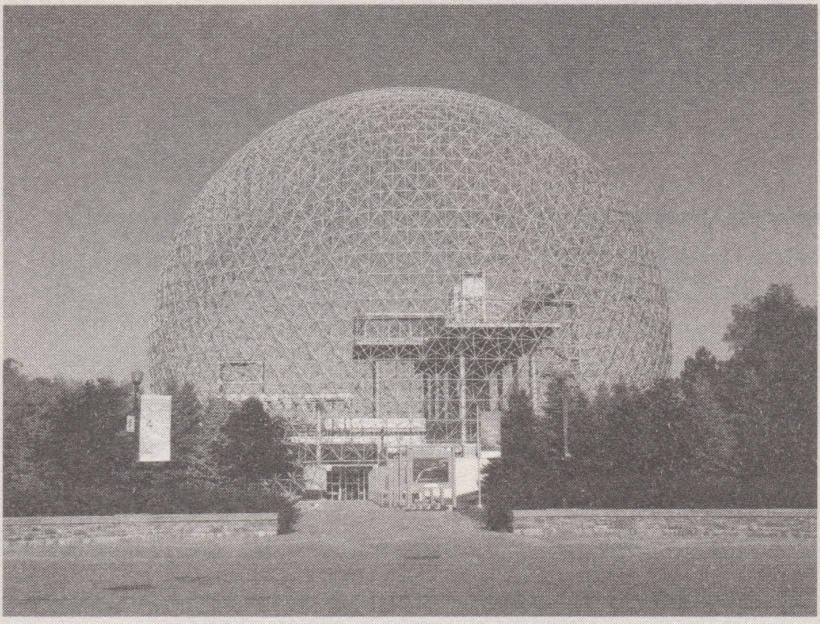
An intuitive proof would be the following. Let's suppose we have two points on a sphere and we want to find the curve that describes the shortest path between them. It seems logical to suppose that we could limit our search to the circumferences of the sphere that go through the two points, which are obtained by intersecting the spherical surface with the planes that pass through those points, as any other curve would be of a longer length than these circles. What's more, by symmetry it can be clearly observed that the arc of the circumference obtained from the plane that passes through the centre of



The arc of the great circle between two points has the least length of all the circle arcs that join those points.

GEODESIC DOMES

One of the 20th century's most surprising and impressive spherical constructions is the geodesic dome designed by Richard Buckminster Fuller (1895–1983). There are so many things we could say about this inventor, architect, engineer, mathematician, poet and cosmologist who was ahead of his time and who strove to put science and technology at the service of society. His greatest success was the geodesic dome, which has a geometrically-charged design.

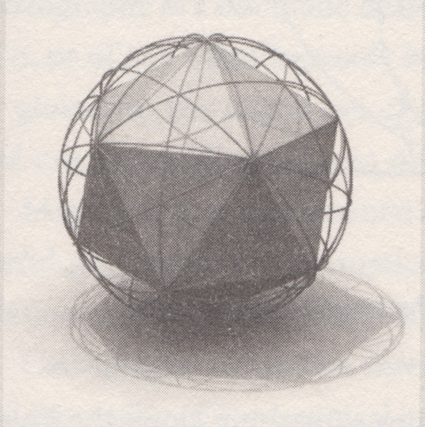


The USA pavilion at the 1967 Universal Exposition held at Montreal, built following a design by Buckminster Fuller. Afterwards the pavilion was converted into a museum dedicated to water and the environment (photo: Philipp Hienstorfer).

the sphere is the arc that has the least length between the points, as is shown in the previous illustration. In short, the great circles are the geodesics of a sphere, that is, the 'tautest' curves, the ones that minimise distances.

The curvature of the great circles

Straight lines are also characterised by the fact that they are curves which don't curve, or, put another way, their curve is null. Does something similar happen with the sphere's great circles? As seems clear from its symmetrical shape, the circumference, which is a planar curve, has the same curvature at all its points and that curvature is not null. What's more, the greater the radius of the circle, the 'tauter' it is and the less its curvature is (see the illustration on the next page). Geometrically, the curvature of a circle of radius r is equal to $1/r$; consequently, the bigger its radius is, the less its curvature. This variation in the curvature of the circumference of different sized

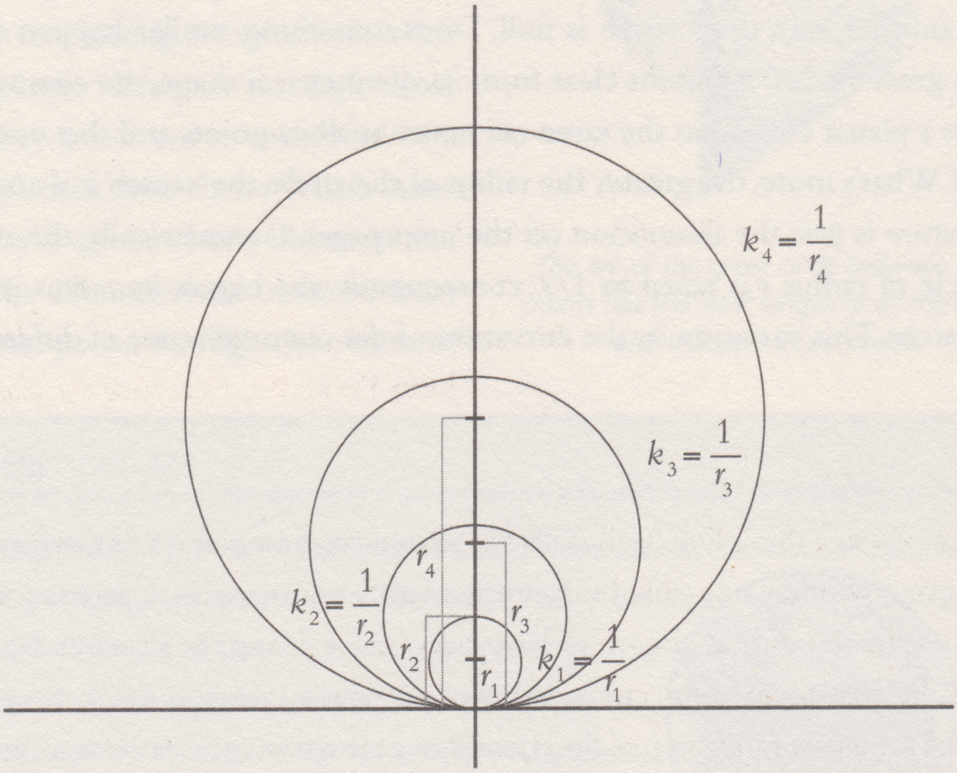


The geodesic dome is a spherical structure composed of a network of great circles (geodesics) which intersect creating triangular elements that give the structure rigidity. The most typical design is to consider an icosahedron inscribed inside a sphere, as in the illustration, and then to divide each one of the faces of the icosahedron in turn into a grid work of triangles and afterwards to project the structure onto the sphere, creating a network of geodesics.

The advantages of the geodesic dome are:

1. It provides a large space without having structural elements in the way.
2. The ratio of volume to exterior surface is optimum, in other words it provides a large area within a minimum surface.
3. It is easy to keep warm in cold zones, as the loss of heat depends on the ratio between volume and exterior surface, which is optimum.
4. It is very rigid on account of its structural strength and its load distribution.
5. It is a light and easily assembled structure.

circles can be perceived when we ride a bicycle round a circular circuit (or, simply, we ride in a circle). Depending on the size of the radius we will have to turn the handlebars to a greater or lesser degree so as to keep within the circle. When we don't turn the handlebars we describe a 'straight path', which is a great circle and whose curve is the least of all. Thus, great circles are those of minimum curvature, and their radius is the maximum.



The bigger the radius r *of the circumference, the smaller its curvature* k.

In fact, geometrists have defined a new curvature, something along the lines of 'the curvature of the curve within a given surface', called a geodesic curvature, and which measures the curvature of a curve within the surface where it exists, i.e. taking that space as the environment space and not the three-dimensional one.

It turns out that the geodesic curvature is precisely zero for the geodesics and, more specifically, for the sphere's great circles, thus generalising the result of the straight lines on the plane.

Chapter 4

In Search of an Accurate Map of the Earth

> *For over 2,000 years the challenge of representing the round earth on a flat surface has posed mathematical, philosophical and geographical problems, which have attracted many kinds of inventors. The use of maps, obviously, dates to before this period. The modern process for drawing up maps, however, got off to a slow start due to the fact that exploration of the Earth as a whole is a relatively recent development.*
>
> John P. Snyder, *Flattening the Earth* (1993)

Cartography as a discipline deals with the graphic representation of the Earth or part of it, or of any other heavenly body, using mainly flat maps but also models in relief or globes. In the digital era, these have become very sophisticated and interactive and can be transformed into new forms of representation.

Maps fulfil two fundamental functions. They are a tool for storing and representing useful geographical information and they help us to understand spatial patterns, relations and the complexity of the world in which we live.

Cartography comprises three fundamental parts. Firstly, the acquisition, analysis and processing of geographical information that will subsequently be used for drawing up the maps. The sources of information are normally direct field observations, which have been the traditional source throughout the history of cartography and continue to be fundamental to it; aerial or satellite information (photographs, radar, remote sensors and so on); and already existing maps, databases and statistical information.

The second part is known by the name 'mathematical cartography' and is concerned with the study of the projections, that is, geometrical and mathematical transformations of the terrestrial surface - which is a curved surface - onto a plane. These are the factors responsible for the way in which the Earth is represented on maps. The term 'mathematical cartography' has a very wide meaning, which we can sum up as the creation and study of a mathematical basis for drawing up maps and for providing answers to theoretical and practical questions concerning related scientific disciplines: the cartography mentioned above, geodesics, geography and navigation, among many others. The branch of mathematics called differential geometry is one of the most important tools in mathematical cartography.

The study of projections has been a central element of cartography. In this chapter we want to take an in-depth look at our knowledge of maps by studying the projections that create them. We will create a classification of projections according to the way they are created, their geometric properties. We will study their characteristic properties, such as the distortion produced by the projections, and look at the fundamental results of mathematical cartography and its implications in real cartography.

The third and last fundamental part of cartography is the design and publication of the maps. Traditionally they have been made from paper, at first hand-drawn and later, with the arrival of the printing press, they could be printed with ever-improving quality. Nowadays new technologies enable maps to be published in innovative and sophisticated formats. What's more, the users themselves are no longer passive receptors but can interact with the map and help to build it.

Another two important facets of cartography are its history and the uses that are made of the maps. A study of the history of maps is fundamental to understanding them, their evolution, their importance in the history of humankind and the differing images of the world that different societies once held and hold today. It is important to remember that understanding the past helps us to understand the future. Lastly, the study of the usage of maps helps us to be much more efficient when re-purposing them, but also contributes to the design of new tools in mapmaking, in particular new projections that solve existing problems or which adjust the design to the requirements of their use.

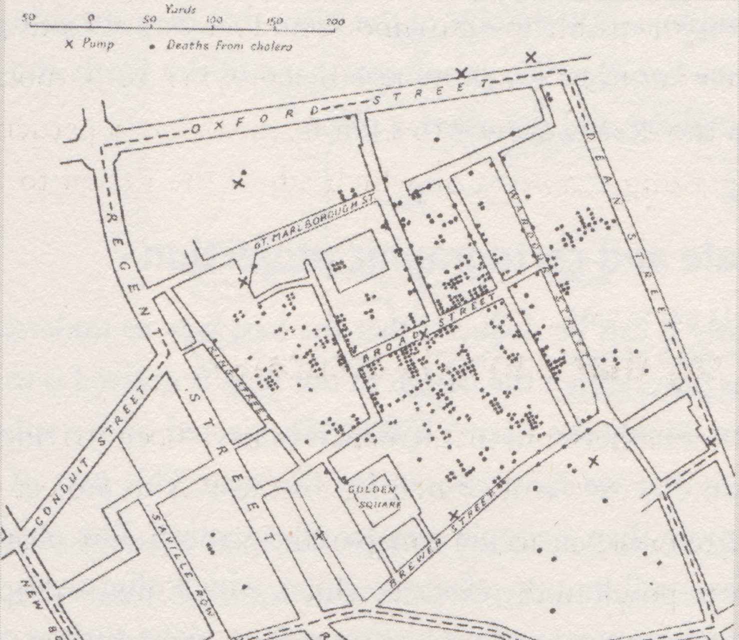
What does it mean to say that a map is accurate?

Throughout history both cartographers and mathematicians have been striving to realise the dream of creating the perfect map, searching for the way to project the terrestrial surface onto a plane to create an accurate map of the planet. In this chapter we shall return to the questions posed in the preface, which can be summed up in one basic question: how do we accurately 'draw' a map of the Earth? But first we should ask another essential question: what does 'accurately' mean?

When we use a map we may have a number of different reasons: to find the shortest route between our starting-point and our destination, and to determine

MAPS FOR SOLVING MYSTERIES

Collecting data of a statistical nature can sometimes help us to make discoveries. One simple example is John Snow's map of the 1854 cholera epidemic. In the mid-19th century, the cause of cholera and other infectious diseases was still not known. The cause was thought to be 'miasmas', noxious substances that travelled through the air. London had suffered many outbreaks of cholera around that time resulting in thousands of deaths. The English doctor John Snow (1813–1858) thought that "water was killing people". At the end of the summer of 1854 there was an outbreak of cholera in London's Soho district. During the first few days more than one hundred people had died; within 10 days more than 500, and in the end, 616. Snow, a witness of one of the 1831 outbreaks, lived in Soho. He suspected that the supply of contaminated water was the possible carrier of the infection. The district's inhabitants got their water from pumps situated in the streets that belonged to two companies that drew the water from the polluted River Thames. Snow drew a map showing the locations of the pumps and the houses where the cholera victims lived, which led him to suspect the cause of the deaths was the water pump in Broad Street around which most of the victims were arranged on the map. He then discovered that all of them had drunk water from that pump. In the end he managed to get it disconnected. Still, years were to go by before it was discovered that water-borne bacteria were the cause of the disease.


Yards
50 0 50 100 150 200
× Pump
• Deaths from cholera
OXFORD STREET
DEAN STREET
REGENT STREET
GT. MARLBOROUGH ST.
BROAD STREET
GOLDEN SQUARE
BREWER STREET
CONDUIT STREET
SAVILLE ROW
NEW BOND STREET
REGENTS QUADRANT
PICCADILLY

John Snow's cholera map showing the cases that occurred in the London outbreak of 1854. The dots indicate the homes of the persons affected by the disease, and the crosses show the water pumps. The dots congregate round the Broad Street pump.

the distance that separates them; to measure the length of a river, a pipeline or lines of communication; to establish the area affected by a missile strike, a gas leak or radioactive fallout; to discover the direction of radio signals or the wind; to set a land, sea, or aerial navigation route; to calculate the surface area of a determined zone such as a forest, lake, urban centre, or agricultural land; to analyse geographical information shown on the map (living standards, pollution, population, economic conditions or the production of different goods, and so on). For this it is essential that the area is maintained on the map, and, if possible, the general appearance of the territories that we are analysing so we can examine and compare the information. That allows us to study geographical features such as hydrographic basins, mountain chains, valleys and coasts, and as such it is very important to preserve their shape and form. In short, we are interested in metric questions such as distances on the map, the length of curves, the shortest distances (the geodesics), the directions, the angles, the areas and the shapes. Therefore, when putting mathematical projections of the Earth's surface onto the plane our aim will be for the projections to preserve those metric features.

Let's pause for a moment and think about the problem of designing a map of the terrestrial surface from an intuitive point of view and without using the tools of differential geometry necessary to provide scientifically rigorous answers. We can think immediately of two difficulties that will appear on attempting a design. One is the difference in scale that exists between people and the surface of the planet that is to be mapped - and consequently the size of the maps that they are going to use. The other is the difference between the geometric shape of the Earth and the map that represents it: one, a sphere, and the other, a plane.

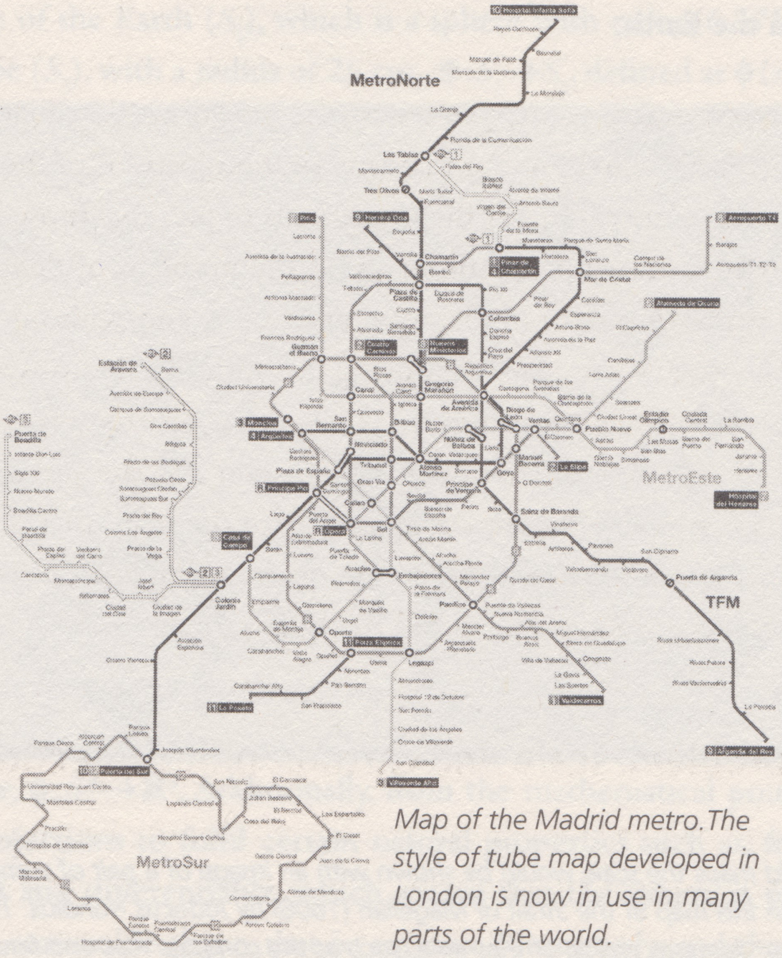
A double problem: scale and cartographic projection

From what was said previously it can be deduced that the best way to understand the mathematical projections from which the design of our map is derived is to see them as a process in two parts. Firstly, the Earth's sphere is projected onto a spherical globe reduced to the size that we have chosen for the map. This part of the projection procedure is no more than a simple process of reduction. The reduced image of the terrestrial surface will be our reference globe, which afterwards will be mathematically transformed by means of projection onto a plane surface, thus creating our map.

The reference globe is a perfect model of the terrestrial surface, the only difference being the size, not the shape. The scale measures the difference between the Earth

TOPOLOGICAL MAPS

In the maps that we might draw of our house, our district or part of the town or city where we live, none of the usual metric issues is preserved. This is what occurred in the case of the first maps made by people, such as the Babylonian tablet from the 6th century BC. They are topological maps in which priority is given to the spatial relations of proximity, separation, order or continuity. They usually represent relationships between places, with the map of a rail network or metro being a good example, as the important thing is not the distance between the stations but to know which stations are connected to which.



Map of the Madrid metro. The style of tube map developed in London is now in use in many parts of the world.

Also included in topological maps are the so-called 'fantasy maps' of imaginary worlds such as the map of the Middle Earth in *The Lord of the Rings* (J.R.R. Tolkien, 1954), or the 'pictorial maps', such as those made for theme parks and zoos. But this type of map also includes maps of neural networks and similar ones used in IT, and those related to graphs.

and the reference globe, and we can obtain it by dividing the reference globe radius by the radius of the Earth. For instance, let's take a 25 cm radius globe representing the Earth, the radius of which can be taken to be 6,371 km (if we use data provided by ellipsoid WGS84); in this case the scale is equal to

$$\frac{25 \text{ cm}}{6{,}371 \text{ km}} = \frac{25 \text{ cm}}{637{,}100{,}000 \text{ cm}} = \frac{1}{25{,}484{,}000}.$$

This scale, which will normally be expressed as 1:25,484,000, means that each centimetre on the globe stands for 25,484,000 cm on the ground – in other words, 254.84 km on the Earth.

Exquisita et magno aliquot mentium periculo lustrata et iam retecta Freti Magellanici Facies
TIERRA DEL FVOGO
MAR DEL NORT
AMERICAE PARS
MAR DEL ZVR
SCALA

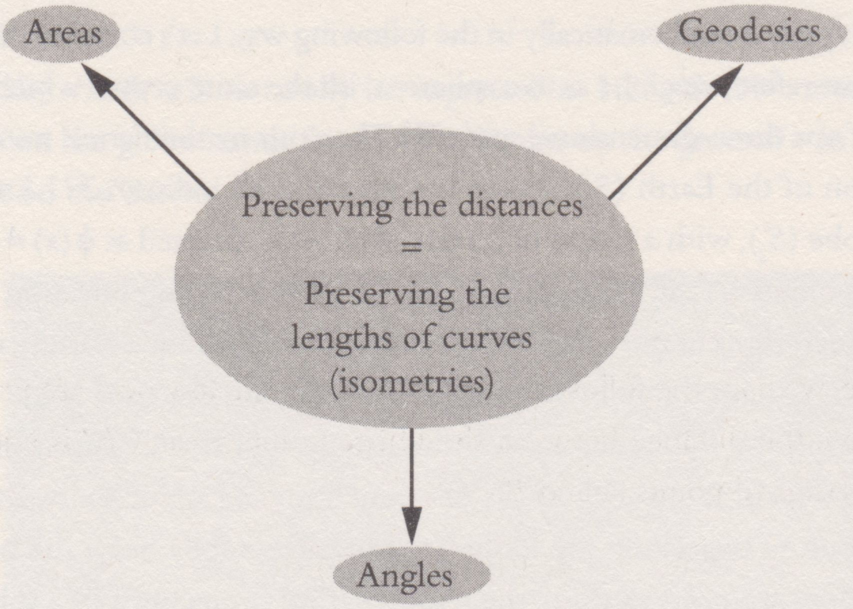
On many old maps the scale would be shown with an image of a pair of compasses, as can be seen on this map of the Strait of Magellan (1606) by Jodocus Hondius. The map also shows other elements typical of that time, such as the compass rose and fantastic animals.

How does this transformation and reduction in size affect the metric issues involved in maps that we mentioned above? The distances and lengths of the curves are modified linearly in accordance with the scale, that is to say, each centimetre represents 254.84 real kilometres. Therefore, if we want to measure the distance between London and Adelaide we simply have to measure it on our reference globe and multiply the centimetres by 254.84 and to give the true distance between both cities in kilometres. The areas are modified in quadratic form, in other words, each square centimetre on the globe is equal to $(254.84)^2 = 64{,}943{,}4256$ km^2. The great circles, which give us the shortest paths, become the great circles on the reference globe, so they preserve the geodesics; the angles - the directions - are also preserved. In short, our reduction of the size of the Earth does not essentially alter the metric issues, except for the scale, which is a constant factor throughout all parts represented on the globe.

We can express it mathematically in the following way. Let's consider the terrestrial sphere and our reference globe as two spheres with the same centre, which we take as the origin of our three-dimensional space $\mathbb{R}^3$. Then our mathematical transformation is a projection of the Earth (S_1), which is a sphere with radius 6,371 km, onto the reference globe (S_2), with a radius of 25 cm., $\phi : S_1 \rightarrow S_2$, defined as $\phi(x) = e \cdot x$, which is what we geometrists call a homothety (if $e > 1$ there is an amplification and if $e < 1$ there is a reduction, as in our case). Mathematically it is a simple transformation and determined in a unique way by its property of reducing the size.

After solving the difference in size, the problem of the variation in shape has to be solved, which, as we shall see, is of a much greater complexity, and is where the true adventure of seeking the Holy Grail of cartography lies – the perfect map. For this purpose we need to study the mathematical projections of a sphere onto a plane – which is the central goal of mathematical cartography and of this chapter – and their behaviour regarding the metric properties. Furthermore, as was pointed out in the Preface, there are many ways of mathematically transforming the sphere onto the plane, and therefore many different projections, creating an equal number of different maps. From here onwards our projections will be applications, point to point, of the sphere of unit radius (with the objective of simplifying the expressions) onto the plane, $\phi : S^2 \rightarrow \mathbb{R}^2$. Additionally, from the mathematical point of view, we expect the projections to fulfil certain natural properties such as, for instance, to be continuous and differentiable, which means that the sphere is projected onto the plane in a reasonable shape, that is, expressed intuitively, and that there are no wrinkles, cut-offs or overlaps.

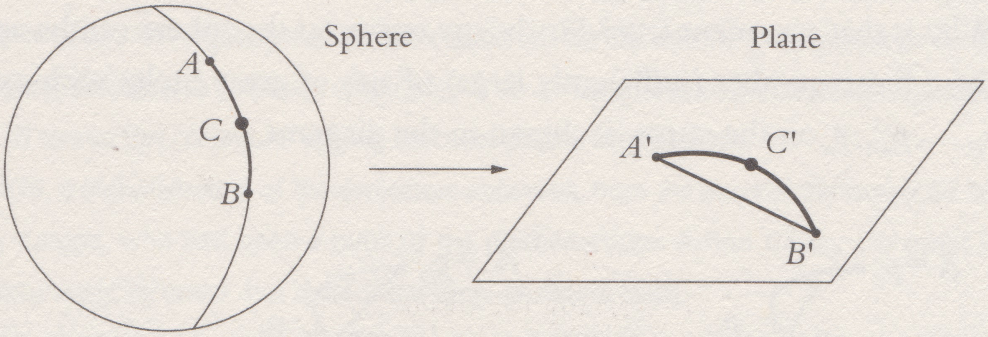
As we mentioned earlier, it is essential to know the behaviour of the projections in relation to metric properties. For that reason, in our search for an accurate map of the terrestrial sphere we will begin by demonstrating that a projection that preserves the distances between points (these applications are called isometries) will also preserve the shortest paths (the geodesics), the angles and the areas. What's more, the property of preserving the distances means that the length of the curves is preserved. These statements are no more than a particular case in the study of the behaviour of differentiable projections between regular surfaces with respect to their metric properties. (A proof carried out with the customary tools of differential geometry can be found in any classical text on that discipline, such as *Differential Geometry of Curves and Surfaces* by M. Do Carmo.)



A projection that preserves distances also preserves the shortest paths

Next we are going to prove that any projection of the sphere onto a plane which preserves distances (this, of course, means that the distance on the sphere between any two points on it is equal to the distance between the images of it on the plane) will also preserve the shortest paths, in other words, the image projected of the great circles on the sphere will be straight lines on the plane.

Let's prove that assertion by *reductio ad absurdum*. As the reader will know, this method of proof consists of supposing that what we want to prove is false and, by reasoning logically, we find a contradiction to the initial hypothesis. Consequently, the result that we really wanted to prove must be true. In the case we are dealing with, we shall suppose that the image of the great circles on the projection is not always a straight line and we shall come up against an incongruity. Let's start on the proof. If, in fact, our projection did not preserve the shortest paths, there would be two points *A* and *B* on the sphere, and one point *C* on the shortest path between them (the great circle that passes through *A* and *B*), in such a way that its image on the plane, let's call it *C*', would not be on the shortest path (the straight line) that joins the images *A*' and *B*', of points *A* and *B*, respectively.



Therefore, we have the following situation. Since the distances are preserved by this projection, the distance between the image points *A*' and *B* ' is equal to that between the original points *A* and *B*:

$$d(A,B) = d(A',B'),$$

and as *C* is on the shortest path between *A* and *B*, the distance between them is the sum of the distances of *A* to *C* and of *C* to *B*:

$$d(A,B) = d(A,C) + d(C,B).$$

However, *C*' is not on the straight line that joins *A*' and *B*'; then:

$$d(A',B') < d(A',C') + d(C',B').$$

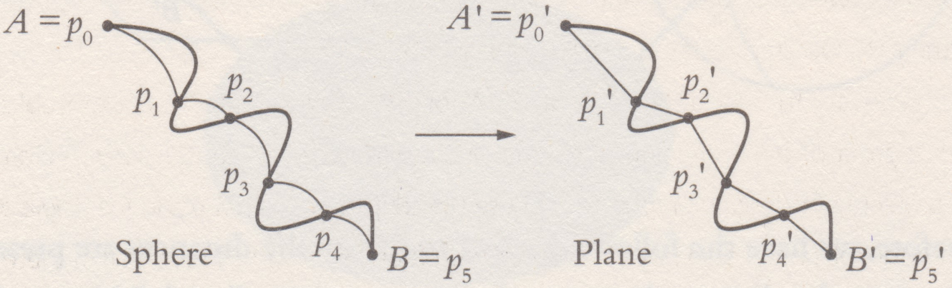
But this latter sum, as the projection preserves the distances, is equal to $d(A,C) + d(C,B)$. Thus we arrive at a contradiction, as we have proved that

$$d(A,B) < d(A,B).$$

The absurdity comes from supposing that the shortest paths are not preserved.

If a projection preserves distances it is equivalent to its preserving the length of the curves

We are going to use the statement from the previous section (the projections that preserve the distances also preserve the shortest paths) to prove that, in such a case, the curves on the sphere are transformed into curves of the same length on the plane. Why is that statement true? Firstly, any curve on the sphere can be approximated by a finite number (sufficiently large) of arcs of great circles with extremes $p_0, p_1, p_2, \dots, p_{n-1}, p_n$ on the curve, as shown in the diagram.



Therefore, we can approximate its length by the sum of the lengths of these arcs or, put another way, by the sum of the distances between its extremes, as the arcs of the great circles are the shortest paths joining them:

$$l(\alpha) = d(p_0, p_1) + d(p_1, p_2) + \dots + d(p_{n-1}, p_n).$$

Secondly, the plane curve, the image of the previous one, is approximated by the straight lines, which are the images of the arcs of the great circles concerned, as we know from the previous section, and the length, by the sum of the distances between the extremes of those segments $p_0', p_1', p_2', \dots, p_n'$:

$$l(\alpha') = d(p_0', p_1') + d(p_1', p_2') + \dots + d(p_{n-1}', p_n').$$

Thirdly, as the projection preserves the distances, then the distance between two consecutive extremes of the original curve on the sphere is equal to the distance of the images, and consecutive extremes of a segment of the image curve:

$$d(p_i, p_{i+1}) = d(p_i', p_{i+1}'), \quad i = 0, \dots, n-1.$$

The three previous statements have enabled us to affirm that the projection transforms a curve on the sphere of a certain length into a curve on the plane of the same length.

HOW MANY COLOURS ARE NEEDED TO COLOUR A MAP?

As children we may have drawn a map in which each region had to be coloured differently from its neighbours. We may have wanted to use more colours, but realised that we only needed to use four. That is what happened, halfway through the 19th century, to the brother of one of Augustus de Morgan's students, Francis Guthrie (who would go on to be a mathematician and botanist), who noted it on a map of the English counties. De Morgan would spread this conjecture among his colleagues in mathematics.

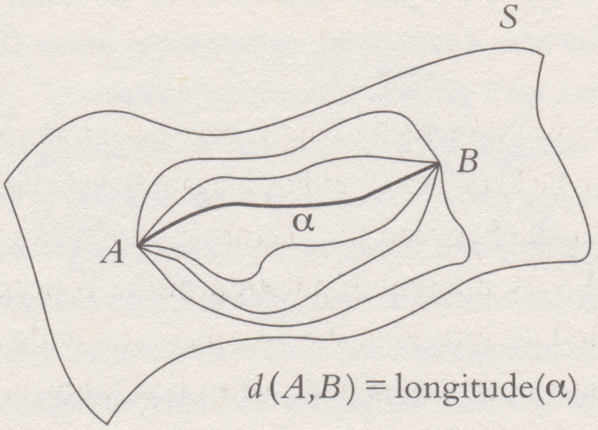
In 1879, a demonstration of the conjecture appeared, from the hand of the lawyer Sir Alfred Bray Kempe, who had been a pupil of the mathematician Arthur Cayley. His proof was, unfortunately, incorrect, but it did show some profound ideas.

It was necessary to wait until 1976 for K. Appel, W. Haken and J. Koch to publish the definitive proof of the four-colour theorem. First it was necessary to transform the problem into a problem of colouring graphs, to suppose the conjecture was false and that there existed maps and graphs that could not be coloured with four colours and so see that in these there were certain 'inevitable configurations' and finally to verify that 'all' these configurations really could be coloured with four hues. The quantity of calculations that had to be made to prove this last step was so huge that it was inevitable that a computer was used. The use of such an instrument in a proof started off a heated controversy in the mathematical community. Can a proof be considered a proof if it includes calculations carried out by computer, when a proof is supposed to be convincing, formalisable and, above all, provable?

A planisphere in four colours, red, blue, green and yellow (here shown in different shades of grey). Four shades are enough for there to be no adjacent regions in the same colour.

In the argument above, a reader with knowledge of differential geometry or familiar with infinitesimal calculus will have noted the point at which, working with more rigour, it is necessary to use the tools of calculus.

The reciprocal result of the preceding result is also true, that is, the projections of the sphere on the plane that preserve the lengths of the curve also preserve the distances between points. This is due to the fact that the distance between two points is the smaller, in fact, the infinitesimal, of the lengths of the curves that join them.

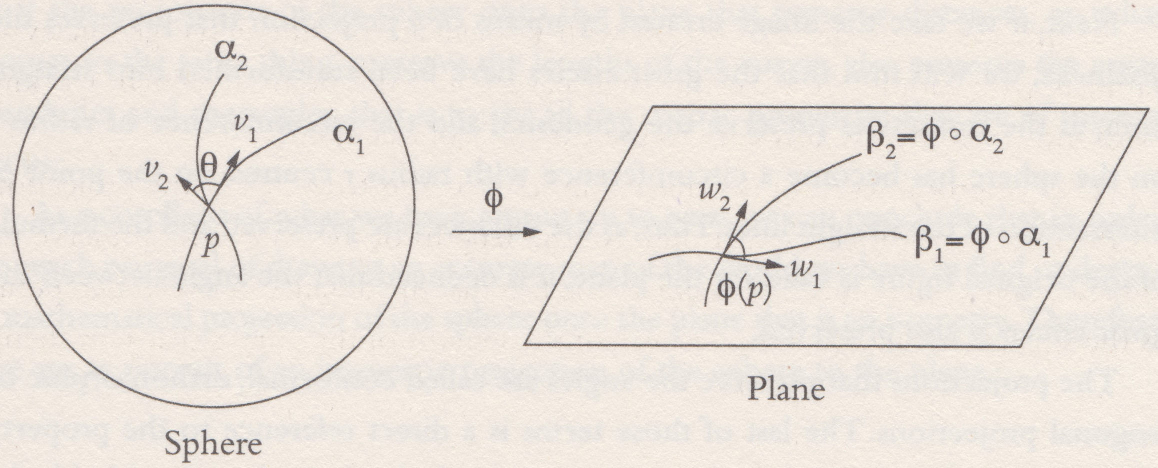


$d(A,B) = \text{longitude}(\alpha)$

A projection that preserves distances also preserves bearings

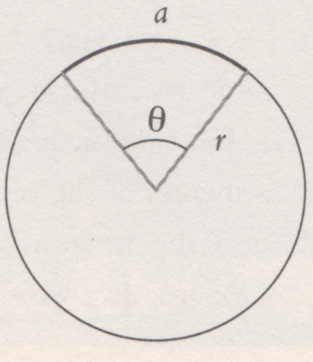
Although we can guess by intuition what it means to say that the projections preserve the bearings – the angles between any two directions – to try to get a better understanding of how the transformations of the sphere onto the plane affect the angles we need to understand the concepts involved a little better, despite the fact that this will involve using some technical terms.

Let's take any point p on the sphere, two directions on that point, that is, two tangent vectors v_1 and v_2, and the angle θ between them. To work out how the tangent vectors are transformed and, consequently, the angle, we'll proceed as follows. We take two curves on the sphere $\alpha_1:(-\varepsilon, \varepsilon) \to S^2$ and $\alpha_2:(-\varepsilon, \varepsilon) \to S^2$ which pass through point p and whose tangent vectors at that point are v_1 and v_2 (mathematically, $\alpha_1'(0) = v_1$, $\alpha_2'(0) = v_2$ and geometrically as shown in the following diagram). Next we turn our attention to the plane curves as images of the previous ones: $\beta_1 = \phi \circ \alpha_1:(-\varepsilon,\varepsilon) \to \mathbb{R}^2$ and $\beta_2 = \phi \circ \alpha_2:(-\varepsilon,\varepsilon) \to \mathbb{R}^2$ and the tangent vectors to them at the intersection point between them, $\phi(p) \in \mathbb{R}^2$, in other words, $w_1 = \beta_1'(0) = (\phi \circ \alpha_1)'(0)$ and $w_2 = \beta_2'(0) = (\phi \circ \alpha_2)'(0)$.



These are the vectors into which v_1 and v_2 have been transformed by means of the projection ϕ. If the angle between w_1 and w_2 is again θ we will say that the projection ϕ has preserved the angle between the vectors v_1 and v_2 (respectively, between the curves α_1 and α_2). A very interesting question of the concepts explained above is that the vectors w_1 and w_2, images of the vectors v_1 and v_2 through projection ϕ, do not depend on the curves α_1 and α_2 that were chosen initially, and so neither does the angle between them. This allows us, for example, to consider as curves α_1 and α_2 the arcs of great circles that pass through point p and with tangent vectors v_1 and v_2, which are unambiguously determined.

We are thus ready to carry out an intuitive proof that the angles are preserved by the isometries. Given two great circles on the sphere that intersect at a point, if we take a circumference of radius r centred on that point, that is to say, formed by the points of the sphere that are separated from the centre by a certain distance r, of a sufficiently small radius, then the angle θ (in degrees) between the two great circles (which is the angle between the two tangent vectors) is approximately the ratio between the length of the arc of circumference determined by the two great circles and 2π times the radius.



$$\theta = a/2\pi r$$

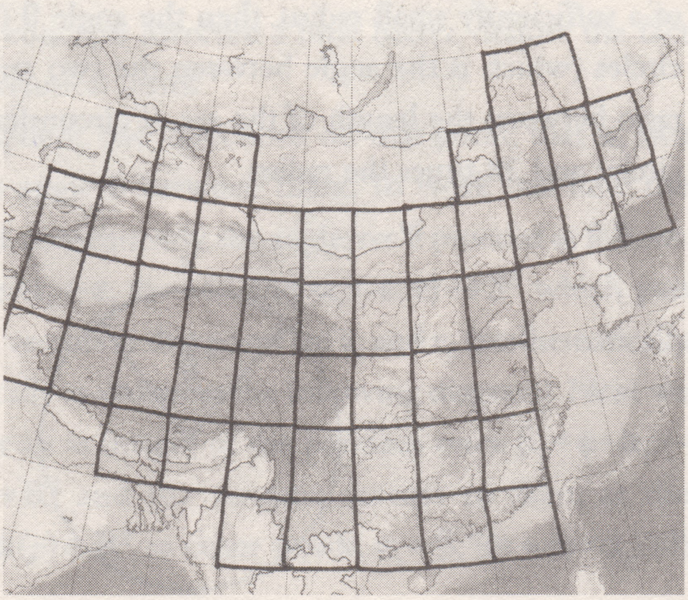
Next, if we take the image created by means of a projection that preserves the distances, we will find that the great circles have been transformed into straight lines, as the isometries preserve the geodesics, and the circumference of radius r on the sphere has become a circumference with radius r centred on the point of intersection of the straight lines. Thus, as the distances are preserved and the formula of the original figure is valid for the plane, it is deduced that the angle between the great circles is also preserved.

The projections that preserve the angles are called conformal, orthomorphic or isogonal projections. The last of those terms is a direct reference to the property of preserving the angles; the other two mean 'with the same shape' and 'with the correct shape'. In actual fact, the projections that preserve the angles preserve the shapes locally for areas that are sufficiently small, but not for those of a certain size. This can be seen, for example, on the Mercator map in the Preface, where there is a very evident distortion that can be seen towards the poles.

A projection that preserves distances also preserves areas

The idea behind the above statement lies in the fact that given a limited area of the sphere, it is possible to cover it with a finite family of regions delineated by meridians and parallels that we can consider rectangular regions (a sufficiently large number of regions which, accordingly, will be sufficiently small), and we can approximate the area of that region by means of the sum of the areas of those 'rectangles' (base × height). On the plane we will obtain the image of the region, covered by the families of image rectangles, and given that the projection preserves the distances, it will have the same area.

The projections that preserve the areas are called equal-area or equivalent projections, or even 'homolographic' projections. Consequently, we have proved



We can approximate the area of a territory, China for example, with the area of a family of rectangular regions delimited by meridians and parallels. The smaller the regions, the more we will approximate the area of the territory.

that the projections of the sphere onto the plane that preserve distances, or, what comes to the same thing, preserve the lengths of the curves, also preserve the areas, geodesics and the angles, that is to say, all the metric properties that are of interest to us.

As a corollary of what we have set out up to now we can conclude that in order to reach our goal of drawing an accurate map of the world we have to find or design a mathematical projection of the sphere onto the plane that is an isometry. Therefore we are in pursuit of an isometric projection of the sphere to the plane.

In search of isometric projection

We have reached the point in this book where we know what we want and, what's more, we know what we need to get it – we are searching for an accurate map of the Earth (or a part of it). Before continuing reading, if we stop for a moment to reflect on what we think might be the way to accurately project the terrestrial sphere onto the plane – in other words, to project it isometrically so that all the metric properties that we are interested in will not be altered – we might assume that the most appropriate way would be to use aerial or satellite photographs and use them to design the corresponding map.

A satellite image of Europe.

However, though it might seem surprising, maps designed from photographs taken from space are not correct maps, because they don't preserve any of the metric concepts analysed above. It's quite a different matter if, for example, we don't care about the margin of error produced on such maps, or if what we want is an attractive image of the Earth seen from space, an image that presents our planet as a heavenly sphere – something that we can sometimes forget when we are dealing with terrestrial maps. The mathematical application that originated from photographing a part of the terrestrial surface, which is the general perspective or scenographic projection, is not an isometry, as the meridians are not represented by straight lines, so the geodesics are not preserved. Neither are the angles preserved, as the images of the meridians and the parallels do not intersect at right angles. Analogously we could analyse the behaviour of the area and find that it is not preserved and, consequently, that neither are the lengths of the curves or the distances preserved.

After this first attempt at creating an accurate map by means of the intuitive scenographic projection, we must continue with our search for isometric projections of the sphere onto the plane. With that goal in mind we could continue creating different cartographic projections, first geometric and then algorithmic, until we find an isometric projection that would allow us to draw up the longed-for perfect map of the Earth. The system is the well-known trial-and-error method, which also has its drawbacks as the number of potential cases may be very great, even infinite, or difficult to identify, and the difficulty increases if the answer is negative.

Instead of starting off by directly creating cartographic projections, studying their properties and ruling them out if they are not isometric, which could be tantamount to looking for a needle in a haystack, we are going to try to limit the field a little. To do so, we'll start by studying if it is enough to find, or create, a projection of the sphere to the plane that preserves, *a priori*, only one of the metric attributes looked at previously, i.e. the angles, the areas, or the geodesics.

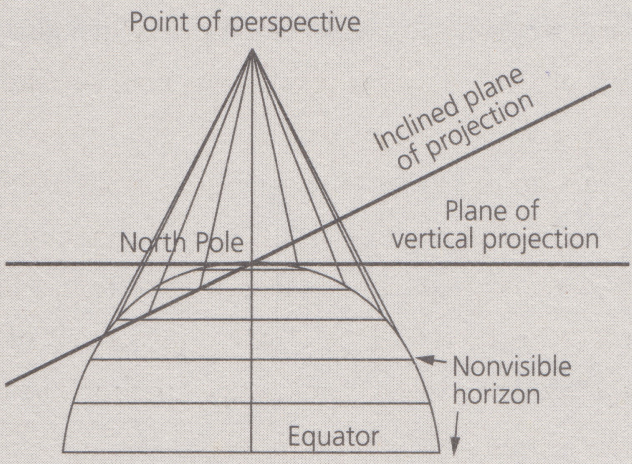
From a formal point of view we would be analysing whether the one-way arrows indicated on the diagram on page 58 are really equivalences (in other words, two-way arrows). Put another way, will it be possible that all the transformations that preserve the angles, i.e., conformal projections, also preserve the distances, the isometries? In the same way, could this happen with the area, or with the geodesics?

From a practical point of view, we would be simplifying our task, as it would allow us to limit our search to the conformal projections (equal-area, or which preserve the geodesics), in other words, we would only have to worry about the cartographic projection preserving one of these geometric attributes.

PROJECTION IN GENERAL PERSPECTIVE

These projections were known to the Greeks and Egyptians more than 2,000 years ago, though they were of little interest until the 18th century, except in non-general cases: orthography, stereographics and gnomonics. If the centre of projection is in line with the centre of the Earth, the projection is known as being 'in vertical perspective'; otherwise it is said to be 'inclined'.

For this projection, the central meridian and parallel are transformed into straight lines, while the other meridians and parallels will be transformed into straight lines, arcs of circumferences or ellipses, and even parabolas and hyperbolae, depending on the aspect of the projection (polar, equatorial or oblique). The metric properties are not preserved, and there is a minor distortion near the centre of projection and an exaggerated deformation towards the edges. This projection, in its normal application, was hardly used except for representing the Earth as seen from space. With the beginning of the space race halfway through the 20th century, it became of interest as it could be used for images of the Earth and other planets. The weather maps we see in the press or on television are normally maps in inclined perspective, as are the simulated images on websites such as Google Earth.



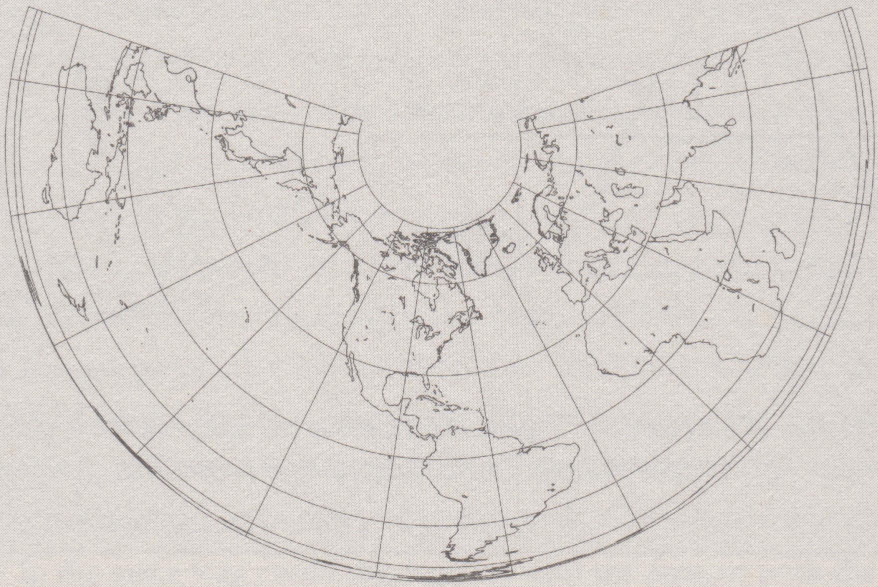
Left, a diagram of the projection in general perspective (vertical or inclined); right, a map created with projection in general vertical perspective.

In the following chapters we will show some concrete examples of conformal projections, equal-area projections and ones that preserve the geodesics, so as to analyse their behaviour regarding other metric attributes and to be able to discover if it is possible to accomplish the criteria mentioned above. On the way, we will be able to look at three historic examples of world maps in which the angles, areas and shortest paths are preserved. In particular, we will learn about the Archimedes projection, which preserves the areas; the central or gnomic projection, which

GEOMETRIC PROJECTIONS VERSUS ALGORITHMIC PROJECTIONS

The fundamental method of classification of cartographic projections is the one based on the method used for creating the projection. By doing this we can classify cartographic projections into geometric and algorithmic (or, likewise, artificial, analytic or mathematical). Geometric projections are those that have a geometric interpretation as 'Sun rays', whether they be understood as emanating from a defined source of light, from the infinite or from a straight line, and that project the Earth (we can think of it, in intuitive terms, as a transparent plastic ball on which the continents and geographic phenomena are drawn), by following the rules of perspective, onto a plane surface or even onto an intermediate surface, such as a cylinder or cone, which is then made into a plane.

The family of geometric projections can be divided into subfamilies if we take into account whether the projection surface is a plane, a cylinder or a cone, which create azimuthal,



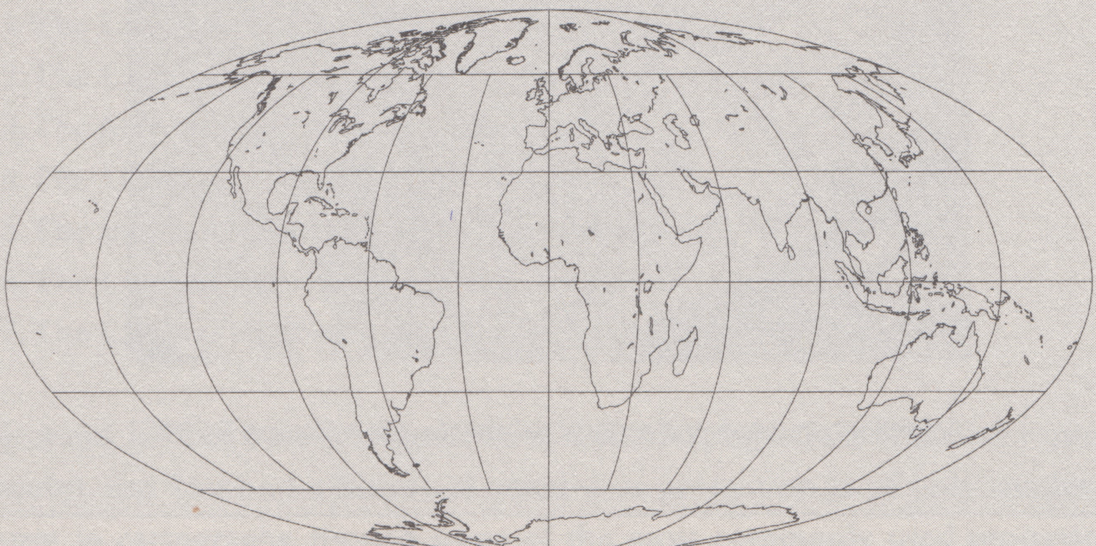
A map created with Alber's equal-area conic projection (1805). It is a geometric projection obtained by projecting the reference sphere onto a cone and then unfolding it into a plane surface.

preserves the geodesics; and the stereographic projection, which preserves the angles. All of these, nevertheless, are not isometric. Consequently, it will not be possible to limit our search just to conformal projections (equal-area ones or those that preserve the geodesics).

cylindrical or conic projections respectively. Examples of geometric projections are the gnomic, the stereographic, Lambert's equal-area cylindrical and Alber's equal-area conic projections.

However, for many cartographic projections there is no direct geometric interpretation as 'Sun rays', and they are simply described by means of mathematical formulae – they are algorithmic projections. Within these, there are projections which indeed are inspired on geometric principles or which are simple derivations of them, such as the Mercator or the Hammer-Aitoff projections. However, there are projections that are purely algorithmic, among them the famous ones by Mollweide, Sanson–Flamsteed, Robinson or Winkel–Tripel.

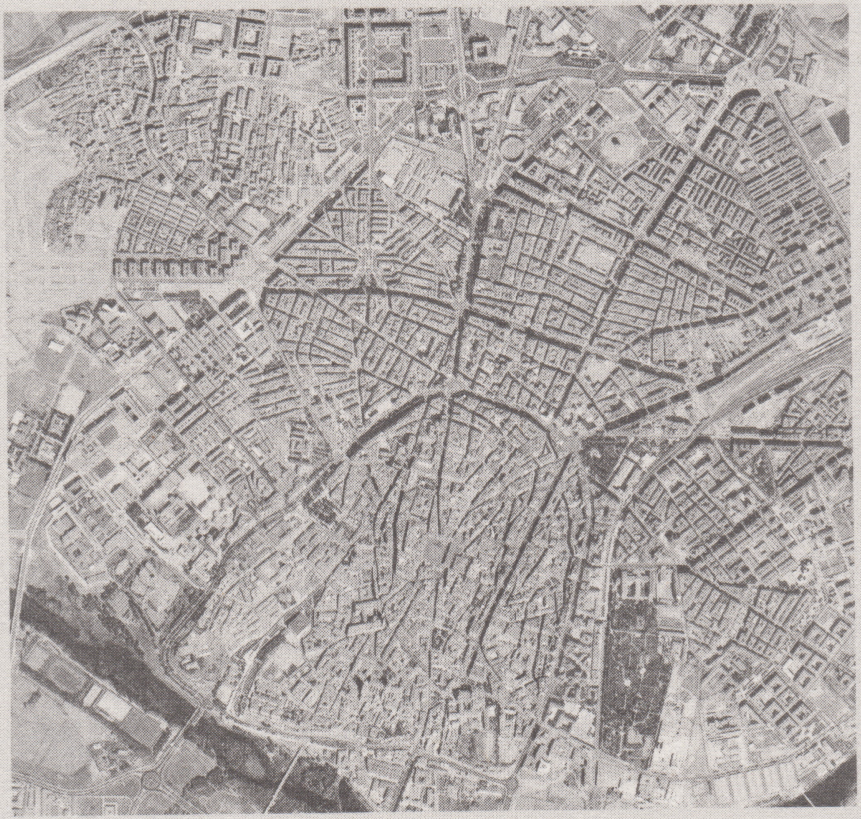
Subclassification of the projections according to the auxiliary surface – plane, cylindrical or conical – continues, to a great degree, in algorithmic projections.



A map produced with a Mollweide projection (1805). It is an algorithmic projection, that is to say, generated from a purely mathematical expression, which is equal-area, in the shape of an ellipse 2:1, and the parallels are parallel straight lines.

ORTHOPHOTOGRAPHY

It is common for tourist maps of small cities or towns to include an orthophotograph. When they see one of these images most people think that it is an exact aerial photograph of the city, in other words, that it is a 'map' in general vertical perspective. This is not completely true. *An orthophoto is a photographic type of representation of a town created by a series of aerial* photos of different parts of it but corrected so that the image corresponds to an orthogonal projection (which is that which is projected through parallel rays, rays with one same direction, and that we can consider a special case of projection in general vertical perspective when the focus is on the infinite), and arranged in a group. It is precisely the use of the orthogonal projection that ensures that there is no distortion on passing from one photo to another.



An orthophotograph of the city of Salamanca, Spain (source: SIGPAC).

Chapter 5

The Archimedes Projection or Lambert Equal-Area Cylindrical Projection

The Lambert equal-area cylindrical projection, also known as the Archimedes projection, due to the fact that the Greek mathematician was probably familiar with this type of map, is one of the seven projections proposed by the mathematician Johann Heinrich Lambert in his work *Notes and Comments on the Composition of Terrestrial and Celestial Maps* (1772). This may well have been the first mathematical work in which general research was carried out on cartographic projections that made use of the new tools of calculus. Following the order in which they appear in it and with the names by which they are known nowadays, the projections that Lambert presented in this work were:

1. The Lambert conformal conic projection.
2. The Lagrange projection.
3. The Mercator transverse projection.
4. The Lambert cylindrical equal-area projection.
5. The transverse cylindrical equal-area projection.
6. The Lambert azimuthal equal-area projection.
7. The Lambert equal-area conic projection.

Projections 1, 3 and 6 are among the most common projections of the last century. However, the Lambert cylindrical equal-area projection is of only moderate interest in cartography, though some of the most commonly used projections derive from it. Its importance is due to its simplicity and rich geometry, which is the reason it is often included in books on cartography as an example of equal-area projections.

Definition and cartographic properties

The Archimedes projection of the sphere onto a cylinder tangential to it is defined as that on which the image of any point *A* on the sphere is the point *A'* on the cylinder which is the intersection of the cylinder with the straight line that passes through *A* and perpendicularly intersects the axis of the cylinder, as the following diagram shows. This projection is clearly geometric. Continuing with the analogy of the Earth as a semi-transparent plastic ball that we used in the previous chapter, the terrestrial image on the cylinder is produced by applying a source of light on the whole of the axis of the cylinder, surrounded by a lens which only allows through those light rays projected horizontally, that is, orthogonal to the axis.

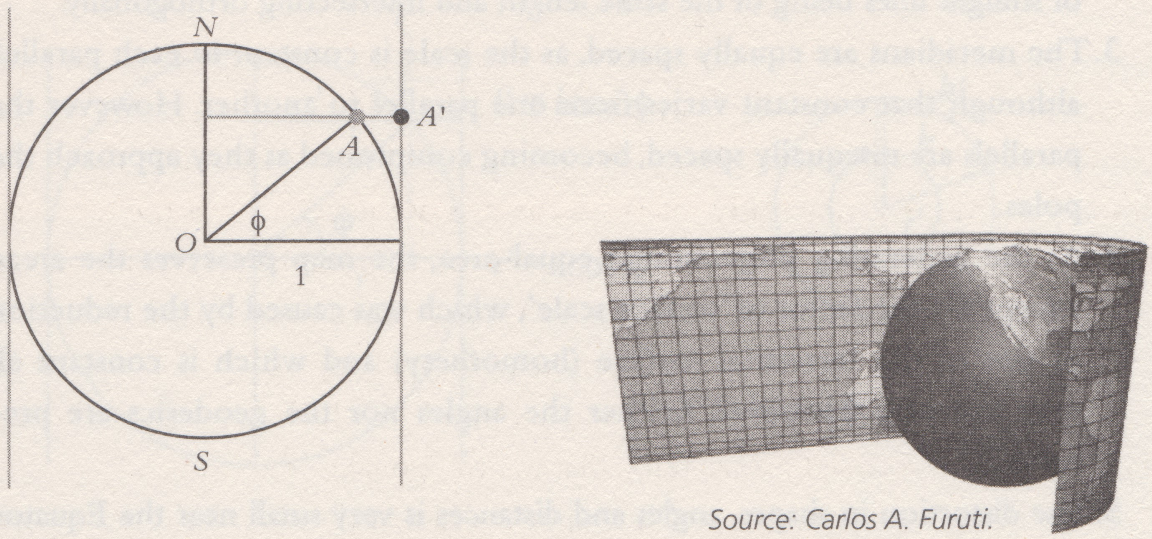
If we consider the Earth as our unit sphere and that the cylinder is the tangent to the sphere at the Equator, then its axis passes through the North and South Poles.

JOHANN HEINRICH LAMBERT (1728–1777)

Lambert was born in Mülhausen (now Mulhouse in Alsace, France), to where his Calvinist family emigrated for religious reasons. At the age of 12 he had to leave school to work with his father, a tailor, though he continued studying on his own in his spare time. Later he worked in metallurgy and as a shop assistant, and in 1746 become the private secretary of the Swiss philosopher Isaak Iselin (1728–1782), in Basle. Two years later he worked as a teacher at the home of Count Peter von Salis, in Chur, a job which allowed him to enjoy free time to study mathematics, astronomy and philosophy as well as providing him with access to the count's excellent library.

Lambert was an exceptional mathematician who proved the irrationality of π and conjectured that the numbers *e* and π were transcendental, in other words, that they are not roots of any polynomial with whole number coefficients. He was one of the first mathematicians to study the problem of Euclid's fifth postulate by assuming it was false and obtaining results in non-Euclidean geometry. He introduced hyperbolic functions, and carried out important research on spherical geometry, cartography and perspective.

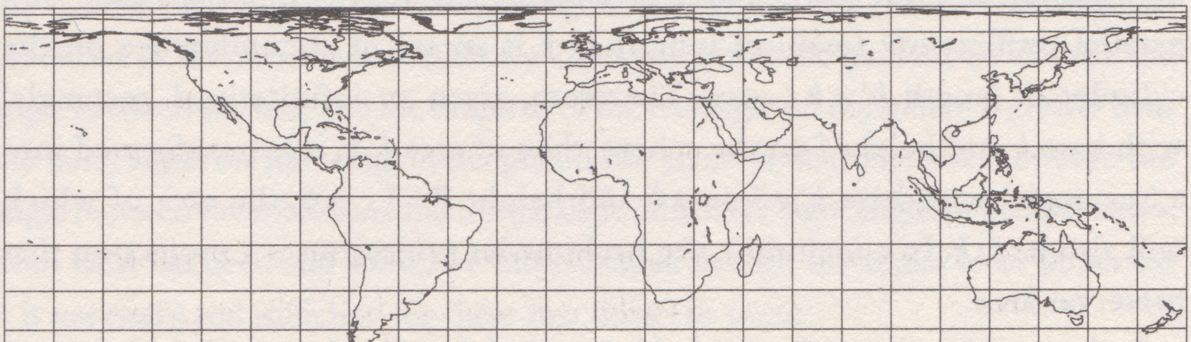
Additionally, he made significant contributions to the theory of probability. Lambert's work wasn't limited to mathematics; he also made important contributions to physics, astronomy and philosophy.



Source: Carlos A. Furuti.

The Archimedes projection horizontally projects the points on the terrestrial sphere onto the cylinder that is tangential to it. The cylinder is then unrolled after cutting it along a meridian.

Once the image of the sphere has been projected onto the cylinder, the cylinder is unfolded onto the plane after cutting along one of its meridians. This unrolling of the cylinder into a plane surface is isometric and preserves all the metric properties. A map of the world was created in this way by J.H. Lambert in 1772.



A map made from a Lambert equal-area cylindrical projection (1772).

These are some of the properties of the map created by using the Lambert equal-area cylindrical projection:

1. It is rectangular in shape, as are all cylindrical projections.
2. The meridians and the parallels are straight lines, with each of the two families

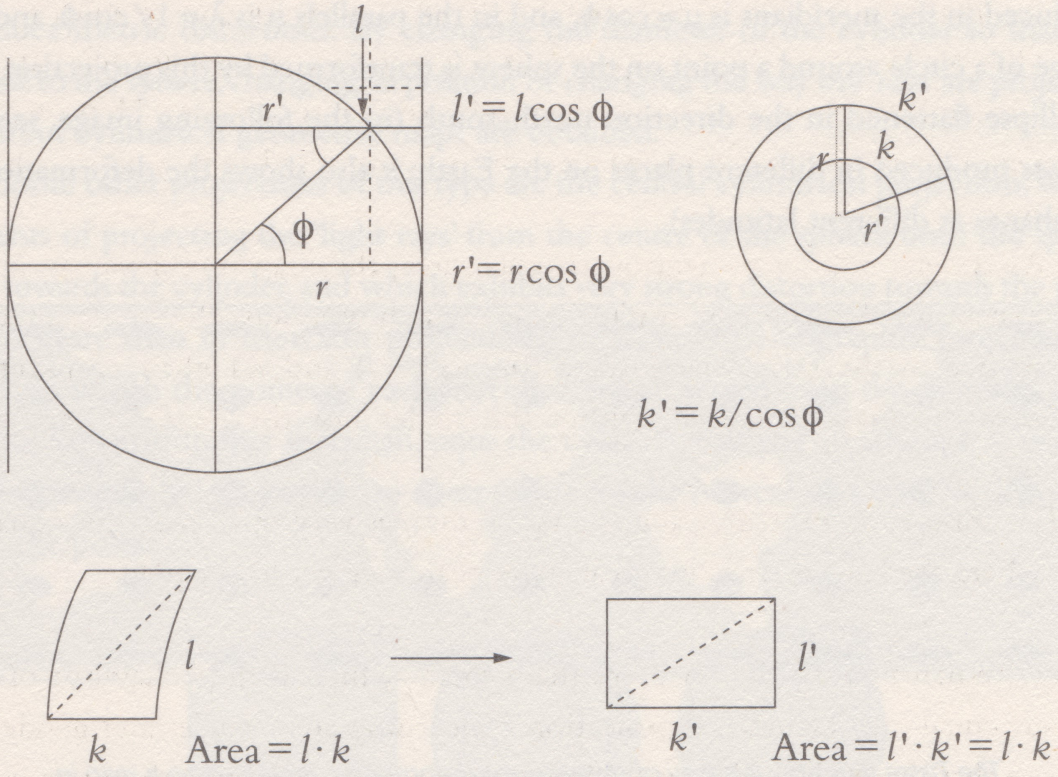
of straight lines being of the same length and intersecting orthogonally.

3. The meridians are equally spaced, as the scale is constant in each parallel, although that constant varies from one parallel to another. However the parallels are unequally spaced, becoming compressed as they approach the poles.
4. As the generating projection is equal-area, the map preserves the areas, except for the factor of 'surface scale', which was caused by the reduction in size of the terrestrial surface (homothety) and which is constant all over the map. However, neither the angles nor the geodesics are preserved.
5. The distortion in shapes, angles and distances is very small near the Equator, where the scale is true, but greater the nearer we get to the poles.

Let's return now to the chapter's main focus, which is the behaviour of the Lambert equal-area cylindrical projection concerning areas, angles and geodesics. To prove that the Archimedes projection preserves areas, it is enough just to prove it for 'rectangular' areas (very small, so-called infinitesimal) the sides of which are meridians and parallels.

As can be seen in the diagram opposite, given a point on the sphere at latitude ϕ, the image of a meridian (sufficiently small) of length l is a segment of a straight line on the cylinder, of length $l' = l \cos \phi$, while the image of a parallel (sufficiently small) of longitude k is an arc of circumference on the cylinder of length $k' = k / \cos \phi$. Therefore, given an infinitesimal 'rectangle' with base k and height l on the sphere, then of area $l \cdot k$, it is transformed into a 'rectangle' with base $k' = k / \cos \phi$ and height $l' = l \cos \phi$, the area of which will also be $l \cdot k$. In conclusion, the Archimedes projection is a projection that preserves area.

Let's remember that, in line with this book's philosophy, this is an intuitive proof, with classical geometry, and that a more rigorous proof would be carried out within differential geometry, making use of calculus.

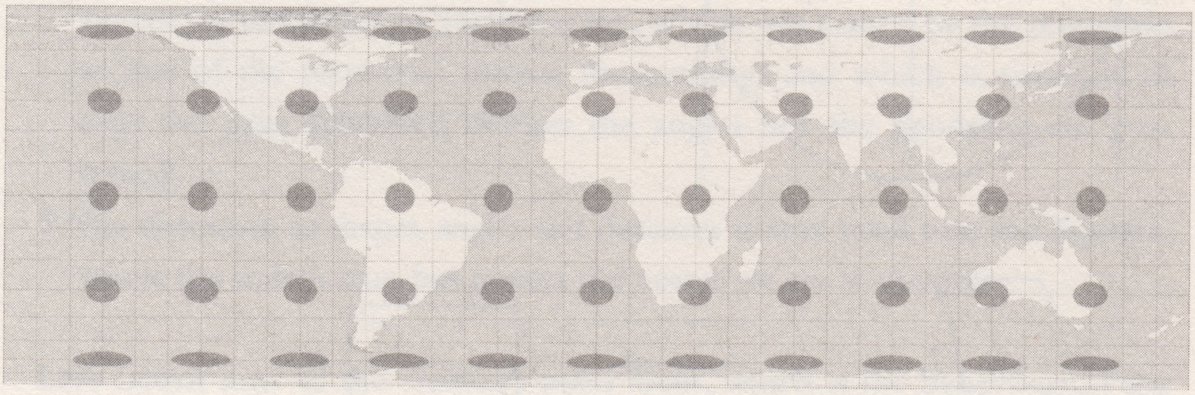


The Archimedes projection is equal-area.

However, on this Lambert map the angles are not preserved. To verify this fact we can use the image of the previous figure. We can see that due to the distortion produced in the meridians (which are contracted) and the parallels (widened), the angle between the base and the diagonal of the rectangle on the sphere is greater than the same angle on the image rectangle on the plane. It does preserve, however, the angle between the meridians and the parallels, which is a right angle. In general terms, from what has been said above, we can conclude that for the angles to be preserved it is necessary and sufficient for these two things to occur:

1. For the angles between the meridians and the parallels (which are right angles or 90° angles) to be preserved.
2. For the distortion in the direction of the meridians μ to be equal to the distortion in the direction of the parallels λ.

In accordance with Pythagoras' theorem, if both properties are fulfilled we can also conclude that the distortion in any direction is always the same. In particular, for the Lambert equal-area cylindrical projection we have seen that the distortion produced in the meridians is $\mu = \cos\phi$, and in the parallels it is $\lambda = 1/\cos\phi$, and the image of a circle around a point on the sphere is transformed by this projection into an ellipse flattened in the direction north-south (in the following image, see the ellipses produced in different places on the Earth; it also shows the deformation in the shapes at different latitudes).



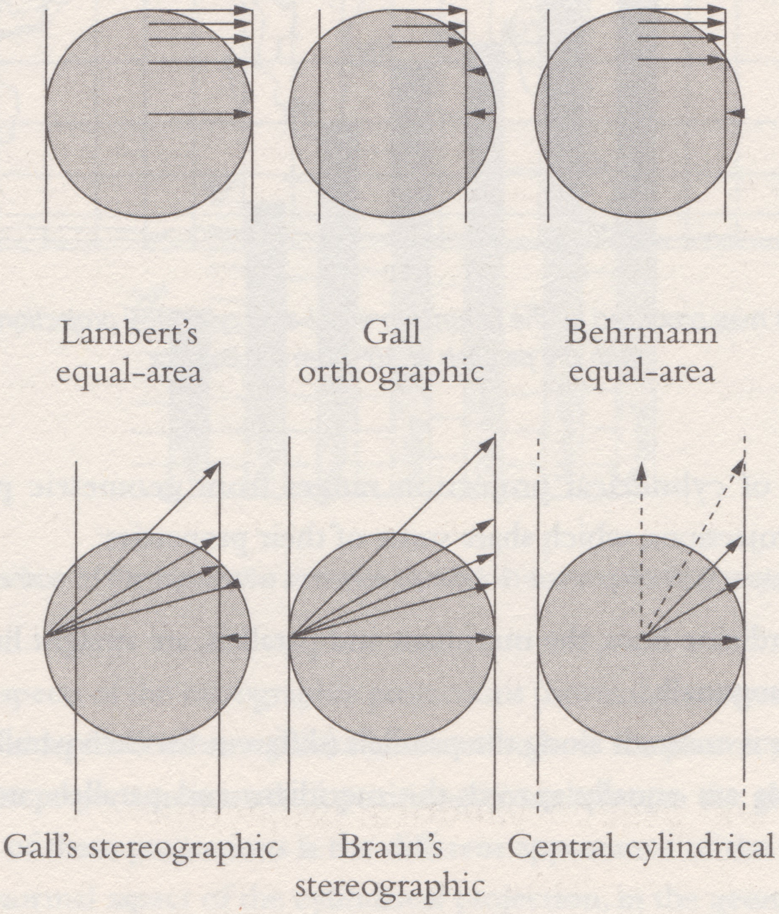
The Tissot indicatrix is a way of showing the deformation produced on a map in graphic form. Small circumferences are taken along the whole of the terrestrial surface and then their images on the map are considered so as to have a representation of the distortion produced by the projection at different places. So, if we consider that the circle has radius 1, the circumference will be transformed into an ellipse of semi-axes λ and μ. If $\lambda = \mu$, the ellipses will be circumferences, and the map will be conformal, and if $\lambda = 1/\mu$, it will be equal-area. Here we see the Tissot indicatrix on a Lambert equal-area cylindrical map.

Finally, it is obvious that this projection does not preserve the geodesics, only the meridians and the Equator. In conclusion, equal-area projections may not be isometric, and the property of preserving areas would not be sufficient by itself to create a perfect map of the Earth.

Cylindrical and pseudocylindrical projections

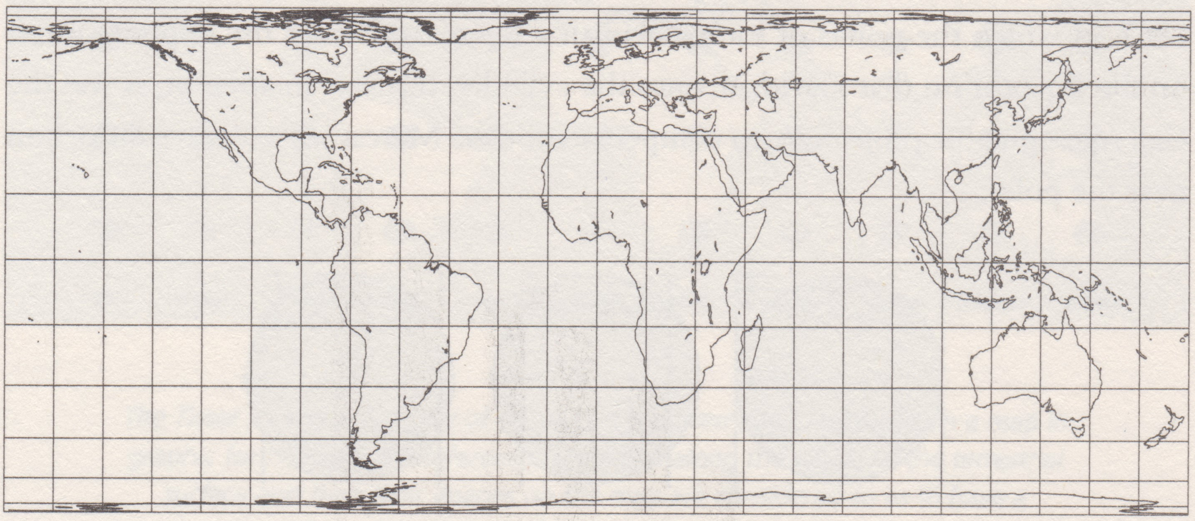
The Lambert equal-area cylindrical projection is a geometric cylindrical projection which is characterised by being created by the terrestrial sphere being geometrically projected onto a cylinder that is a tangent, at the Equator in the normal versions, to the Earth, with the cylinder afterwards being unrolled into a planar surface, which is normally done by cutting the cylinder along one of the meridians, that is, one of its vertical straight lines. The distortion on this type of map is produced in the first stage, as the unrolling of the cylinder onto the plane is isometric and does not produce metric distortions. By changing the diameter of the cylinder so that it is secant to the sphere, changing its position or changing the way the 'rays' are projected, different cylindrical geometric maps are obtained.

Some other projections of this type are the central cylindrical projection, which consists of projecting the 'light rays' from the centre of the sphere onto the sphere and towards the cylinder, and which exhibits very strong distortion towards the poles (even more than in Mercator projections) or Braun's stereographic projection of 1867, in which the points of each meridian are projected from the opposite point on the Equator on that meridian, onto the cylinder; this was an attempt, as was the Gall stereographic projection, to compensate for the Mercator's excessive distortion near the poles.



A diagram of a section of some geometric cylindrical projections which show the differences in size and aspect of the maps they produce.

We are considering the case when the cylinder is tangent to the sphere at the Equator, but we should also bear in mind the cases in which the cylinder is secant to the globe, intersecting it through two symmetrical parallels with respect to the Equator. So, the same geometrical projection for the Lambert equal-area cylindrical projection will create, for example, the Behrmann projection (1910), if the intersecting parallels are at 30° N and S, or the Gall-Peters projection (1855 and 1967) if they are at 45°, both of them being equivalent. The secant version at 45° of the Braun stereographic projection is the Gall stereographic projection (1885).



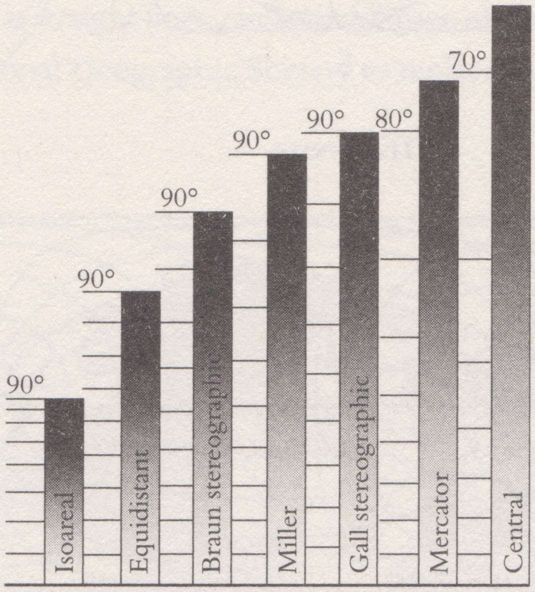
A map produced by the Behrmann equal-area cylindrical projection, with the parallels at 30° from the Equator.

The concept of cylindrical projection ranges from geometric projections to algorithmic projections, which share some of their properties:

1. The coordinate lines, the meridians and parallels, are straight lines that intersect orthogonally.
2. The scale is constant along the parallels (different for each parallel), and so the meridians are equally spaced; the meridians and parallels are of the same length.

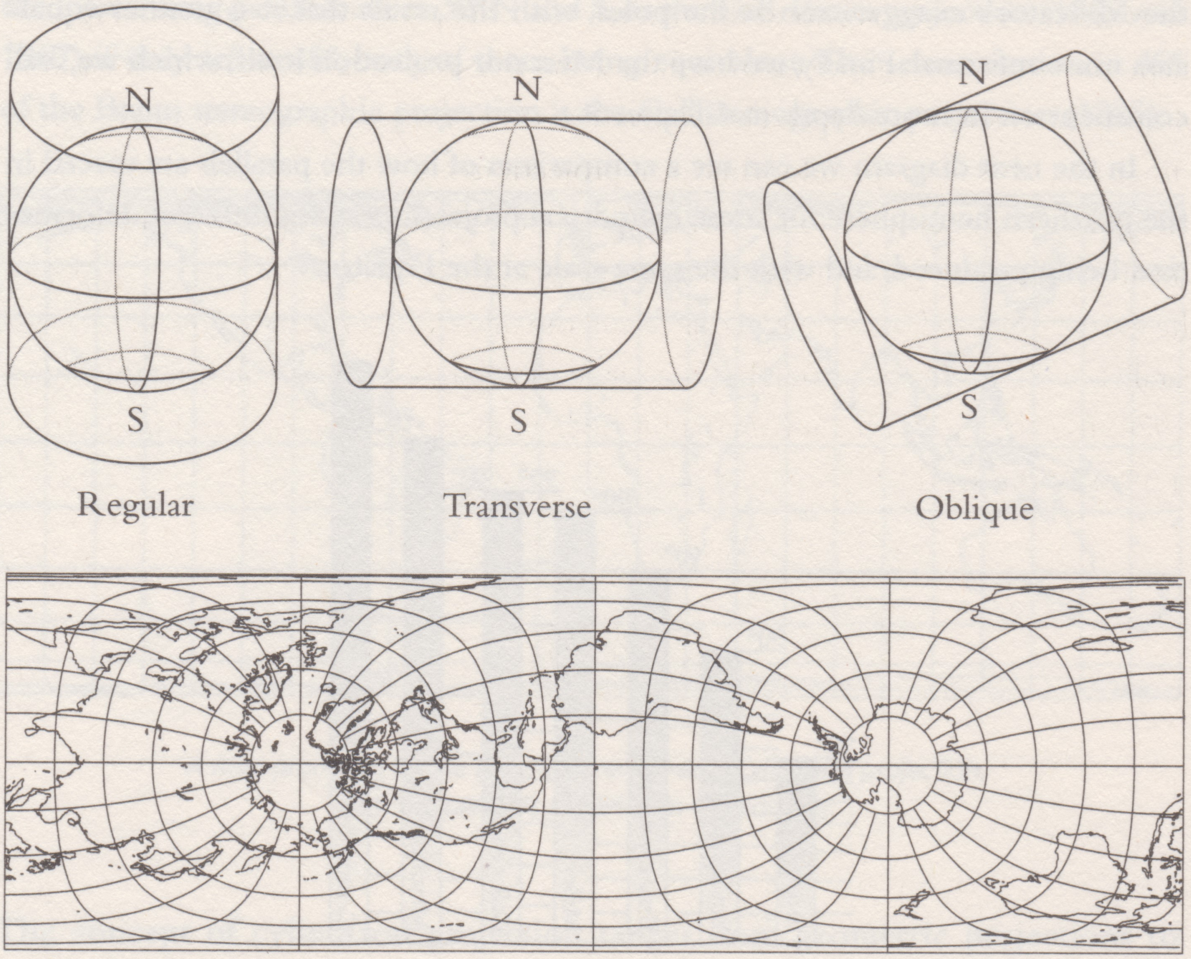
In particular, the world maps produced by these projections are rectangular and show symmetrical behaviour with respect to the Equator. Among the examples is the equidistant cylindrical projection. In this simple projection, which was introduced by Eratosthenes, the meridians also have constant scale resulting in the parallels being equally spaced, a particular case being when the separation between meridians and parallels is the same, forming a graticule known as Plate Carrée, and which is the mathematical projection with the simplest formulation as on the plane it simply represents latitude ϕ and longitude θ. Miller's cylindrical projection (1942) was an attempt to keep the shapes of the Mercator projection, but reduce the Mercator's exaggeration at the poles, with the result that it is neither equal-area nor conformal. Finally, we have the Mercator projection itself, which we shall comment on in more depth in Chapter 9.

In the next diagram we can see a comparison of how the parallels are spaced in the northern hemisphere for some cylindrical projections, with different deformation being produced, and with the same scale at the Equator.



Comparison of the separation between parallels in some cylindrical projections.

Different aspects of the cartographic projections (normal, transverse and oblique) can also be considered. These are generally related to how the plane, cylinder or cone onto which the projection is made is positioned in relation to the globe. The main consequences of these projections is the different appearance of the meridians and parallels. The normal aspect of the cylindrical projection, in the geometric case and by extension also in the algorithmic, is that in which the cylinder is tangential to the sphere at the Equator or secant on a pair of parallels, which is the case mentioned above. The transverse aspect is that in which the cylinder is tangential to a meridian, or secant in circles parallel to the meridian, while we refer to the aspect being oblique when the tangency is a great circle which is neither a meridian nor the Equator, or the intersection curves on the cylinder and sphere are circles parallel to one of these great circles. The fitness of using the transverse or oblique aspect lies, for instance, in the need to focus attention on a region stretching along a meridian, in respect to a generic great circle, as the distortion is lesser around the areas of the tangent lines.



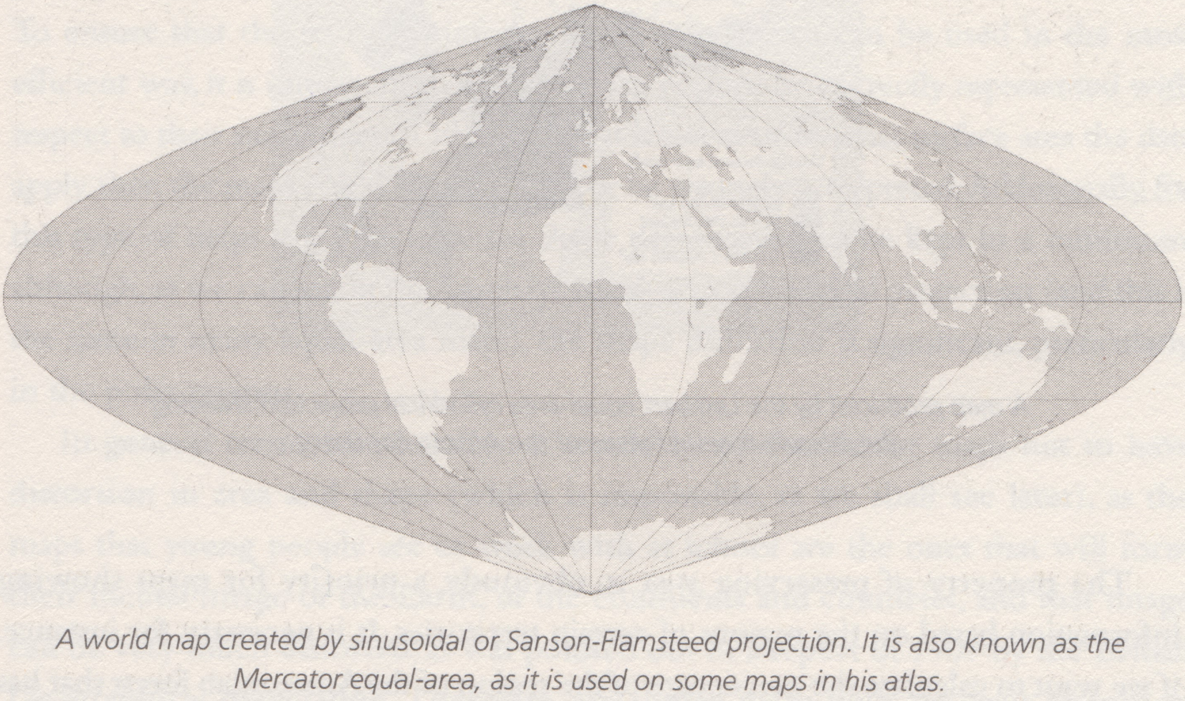
The diagrams above show different aspects of cylindrical projections.
The lower image is the Lambert equal-area cylindrical projection (1772), the tangential meridian of which is 90°W, crossing North America from north to south.

In their normal version, cylindrical projections produce strong distortions in shape, and very often in area too, in regions close to the poles. Furthermore, the world maps they produce can create an unreal image of the world in the minds of some people, particularly schoolchildren.

Pseudocylindrical projections are an attempt to provide a solution to both problems by compressing the parallels as they approach the poles. In the normal aspect, tangential to the Equator, pseudocylindrical projections exhibit the following characteristics:

1. The parallels are horizontal straight lines that are not necessarily equidistant.
2. The meridians are arbitrary equidistant curves along all the parallels.

Therefore, as in the cylindrical projections, the scale is constant along the parallels; however, as the meridians and the parallels do not intersect orthogonally, none of these projections can be conformal. Maps of this type commonly used in atlases are two equal-area projections: the Mollweide (1805) a global map in an elliptical shape on which the meridians are ellipses; and the sinusoidal or Sanson-Flamsteed projection (possibly introduced by Mercator), whose meridians are sinusoidal curves. Other maps based on pseudocylindrical projections are the Collignon (1865) in its triangular or diamond versions, with straight-line inclined meridians, which preserve the areas but show great distortion in shape; the series of six Eckert maps (1906) – the even-numbered ones are equal-area and the odd-numbered have equally-spaced parallels with meridians as straight lines, circles or ellipses; or the Robinson projection (1974) used by the National Geographic Society to replace the Mercator projection in maps of the world.

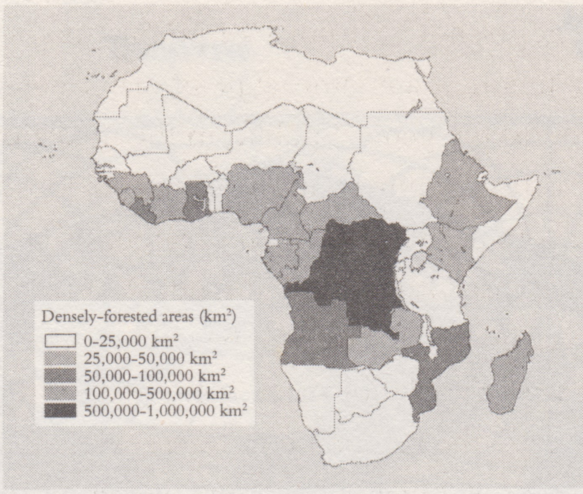


A world map created by sinusoidal or Sanson-Flamsteed projection. It is also known as the Mercator equal-area, as it is used on some maps in his atlas.

The use of equal-area maps

On the use of the Lambert equal-area cylindrical projection, we mentioned earlier that it is of only moderate interest in the designing of maps covering the whole of the terrestrial surface, due to the enormous distortion it creates in the polar regions, except in books on cartography or differential geometry as a simple example of making a map that preserves areas. There is more interest in some of the variations of the normal version of this projection, such as the Gall–Peters (which we shall look at in detail in Chapter 9) or the Behrmann, which is often used in the design of global maps.

However, the moderate distortion produced around the line of tangency makes it very appropriate, in its normal version, for maps of the tropical regions, or for world maps focusing on that area, such as maps depicting the world production of rubber, wood, rice, sugar cane or other products that are particularly tropical in nature. In its transversal version it is useful for maps of regions stretching north-south, while for zones along a great circle the aspect would be oblique. The latter has been used for maps of Eurasia and Africa, and for aviation routes.



A map produced by the Lambert equal-area cylindrical projection showing the densely-forested areas of the African continent .

The property of preserving area is, obviously, a priority for maps showing information based on the regions of certain territories. It is necessary, for instance, if we want to calculate the percentage of the region of the Amazonian forest that has been destroyed, or if we are planning an agricultural division of a province or region, or if we need a map showing the number of inhabitants per square kilometre. In a map of this type the area is to scale: 'surface scale'. (Remember that it should not be confused with normal scale, i.e. the scale of distance, which is linear, and which, for example, on the Lambert equal-area cylindrical map is not constant throughout the whole the map, as can easily be seen from the fact that the length of all the parallels on the map is the same whereas it is not so on the Earth's surface.) We can even use surface scale to calculate the areas of territories simply by working out the area of the zone on the map and multiplying it by the relevant scale.

However, apart from some special cases, maps are not consulted with the aim of calculating area, since such information, when it refers to a particular territory, can be expressed in numerical form. Instead, the usefulness of these maps is that they enable us to contrast the areas of different territories, to compare the surface area of a certain region with that of its neighbours or other regions in the study, or to compare at a glance an area of a region in respect to the whole, whether it be a country, continent or the whole planet.

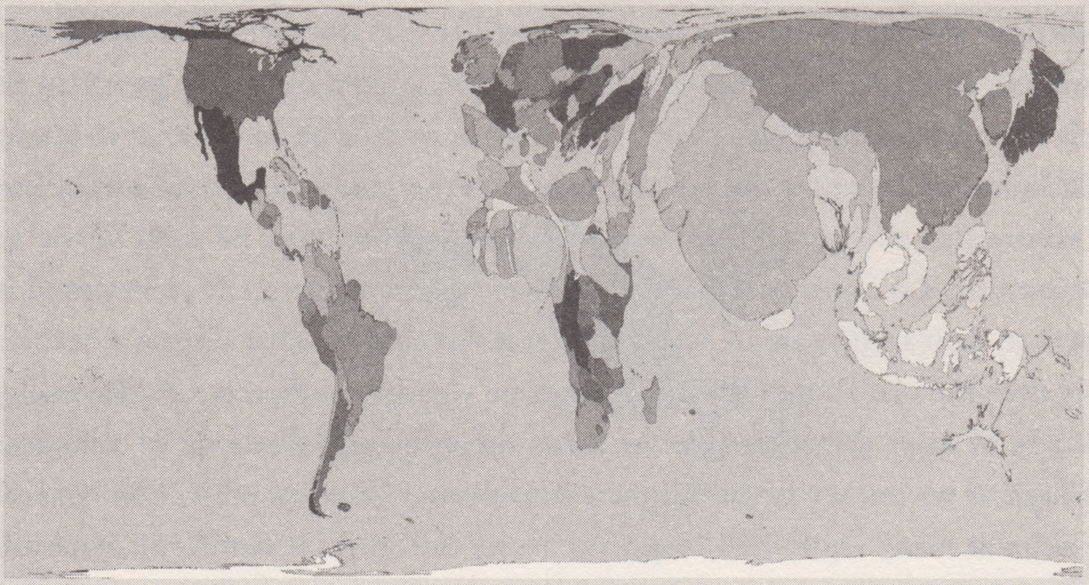
And as maps are an instrument for transmitting information in a faster, more illustrative and meaningful way than a table of numbers, it is important that equivalent projections should be used for maps for popular use, whether for showing species of animals or vegetation, schools, airports, or production of certain goods, to show levels of pollution, religion or any other type of map to be used either in scientific publications or those for the general public, in the media or in educational books. To ensure that the statistical information being shown can be used in the most efficient way, it is important that the relevant regions be correctly represented with respect to their surface area. Otherwise, if it is not clear to what surface area the data apply, then the maps could lead to errors or be wrongly interpreted. Additionally, for this type of maps it is preferable for shape deformation to be kept to a minimum, although, as we can see in the Lambert equal-area cylindrical projection (and this is the same in many equal-area maps), the shape distortion is significant, particularly in the polar regions.

In general terms, in education it would be preferable for maps not to have distortion in area and shape (which is impossible, as we shall see later), as the maps that young people see or work with at school are the ones that will form their mental image of the Earth, of the continents and countries, and that image can be very much distorted, as was pointed out in a report in 1907 by the British Geographical Association. Therefore, equal-area maps with the least distortion possible should be used, or those in which area and shape distortion is slight, such as the maps employed by the National Geographic Society, the Robinson and the Winkel-Tripel.

Some other equal-area projections are the Alber's equal-area conic (1805), the Mollweide (1805), the Gall-Peters orthographic (1855 and 1967), the Eckert IV (1906), the Lambert equal-area azimuthal (1772) and the sinusoidal (or Sanson-Flamsteed, 1650 and 1729) projection, among many others.

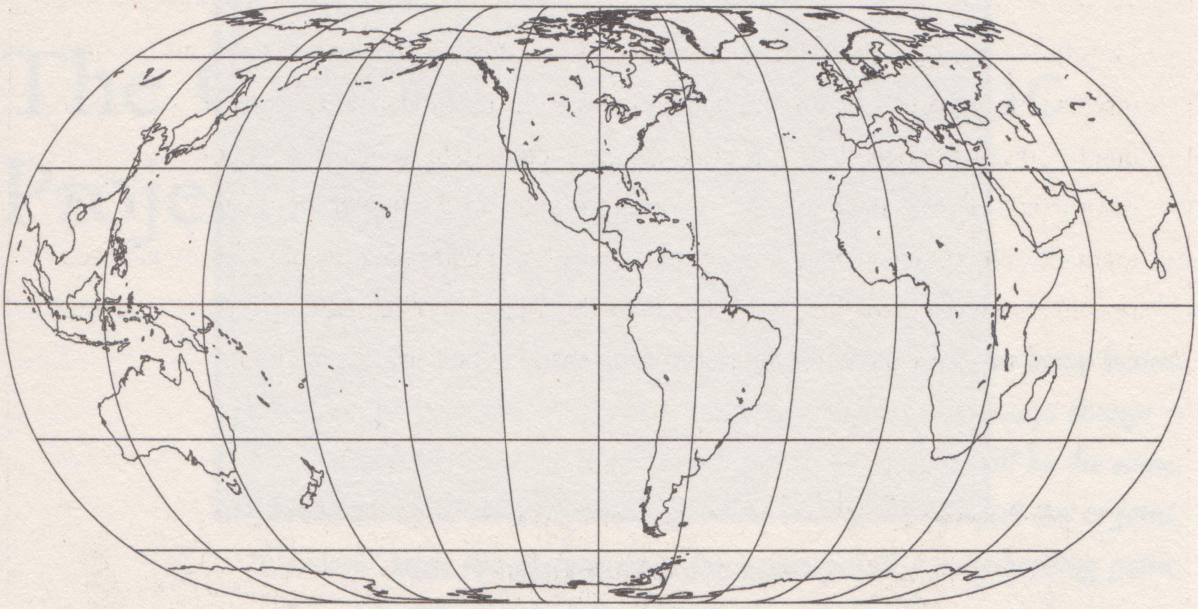
CARTOGRAMS

Cartograms are maps on which the surface of countries or the different regions appears rescaled in proportion to the population or some other similar property, such as birth or death rates, incidence of certain diseases, production data, levels of pollution, economic parameters, poverty levels or some other information on the human population. Although there are different techniques and methodologies, cartograms are normally expected to keep the different regions adequately located in relation to one another (although non-contiguous cartograms also exist) and for the distortion in shape to be kept to a minimum.



A cartogram showing the distribution of the world population by country. Its surface area is determined by the corresponding population distribution (© Copyright SASI Group. University of Sheffield).

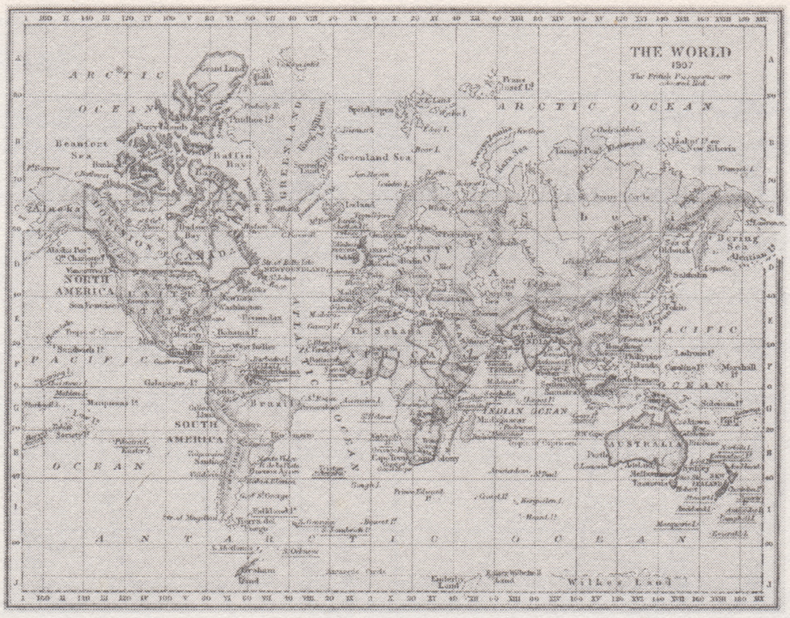
One of the advantages of using cartograms is that they enable the relevant information on different land areas to be compared rapidly. The impact on the reader is usually powerful and impressive but the technique does have some drawbacks. The geographic reality is lost, they can be difficult to 'read' and they give the impression of being incomplete, perhaps because of the geographical uprooting. It is a good idea to place them alongside equal-area maps. In any case, they are useful for attracting the reader's attention and encouraging a critical perspective.



In 1906, the German geographer Max Eckert published six pseudocylindrical projections; those which are denoted by an even number are equal-area ones, such as the projection shown here, the Eckert IV.

Although it seems that things are gradually beginning to change, it is, unfortunately, still common for insufficient care to be taken in the choice of maps for use in education, for scientific purposes, for popular consumption and in the media. That said, however, sometimes the conscious use of maps that do not preserve area can be due to attempts at manipulation for political reasons. They may, for instance, be used for political propaganda so that a region or country looks bigger or more important, with the object of demonstrating its political, economic or military power, and of intimidating the people of other territories. Many countries took advantage of the dominance of the Mercator map, due to its significance in early exploration, to design biased world maps. The British made use of it, and normally the British Empire appeared in red, much more powerful than the other colours used, with the Greenwich meridian and, therefore, London, in the centre, and its colonies in regions of the map where the size of surface area was accentuated. In some cases Australia and New Zealand appeared twice, on both sides of the map!

To demonstrate their power to the rest of the world, Imperial Russia and, later, the communist Soviet Union, took advantage of the fact that the Mercator global map made Russia look enormous. But also, in the opposite way, some political groups raising awareness about the supposed 'red menace' – the spread of communism around



A map of the British Empire in 1907.

the world – also used this map to show the supposed excessive territories of the USSR and China in an aggressive intense red. Likewise, both Europe and the United States were happy with a map that was generous with their image as shown to the world or at least, in general terms, they felt no need to change it.

On the other hand, a country may be shown smaller so as to attempt to prove that it is a victim of aggression by other countries or territories. Thus, in the book *How to Lie with Maps*, the author gives the example of a map published in 1973 by the Jewish National Fund of Canada which, to demonstrate the 'Arab lie' about Israeli aggression, showing Israel in white in the midst of an Arab zone, shown in black, running along the north of Africa, stretching from Morocco to Saudi Arabia. The map spoke about the extension of certain territories, but made no reference to technology, military power, economic power or international support. The Nazis, incidentally, who were experts in the use of political propaganda and particularly of maps, had already used the same argument with a map entitled "Germany, the aggressors?", comparing German territory with that of the whole of the British Empire. These are just a few examples, but the history of cartography includes many others.

Chapter 6

The Central or Gnomonic Projection

For ever Einstein will be tough to tackle with our puny brains.
Whatever be your shutter-speed – and shutter-speed can always change –
whichever way the shadows fall, the speed of light will be the same.
There are other words and concepts here that make you feel out of joint.
'Geodesic' takes the shortest route from one spot to a neighbouring point.
On this complex chess-board, Sir, see where your coordinates are;
to put to test through mathematics your theories of near and far.
Keki N. Daruwalla, 'Space-time instruction'
The Map-Maker (2002)

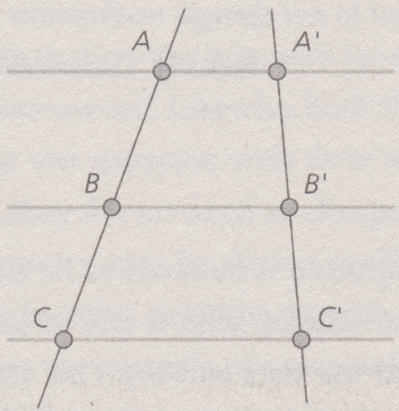
The central or gnomonic projection is believed to be the oldest projection and is usually attributed to Thales of Miletus, who is said to have used it for his oblique view when designing maps of the stars and sky. This azimuthal projection of the terrestrial globe to a plane tangential to it was formerly called *horologium* or horoscope, on account of its connection with sundials, which are fitted with gnomons. On sundials the gnomon is inclined and pointing towards the North Pole, and its shadow tells us the time throughout the day as the Sun crosses the sky. The angles between the hourly marks on a sundial designed for a particular latitude are equal to the angles between the meridians on the gnomonic projection focussed on that same latitude, by counting 15° of longitude from the central meridian as one hour extra from the solar noon.

The origin of the use of the term 'central projection' is not known, while the term 'gnomic projection' was coined in 1749 by the English mathematician William Emerson. In 1836, British mathematician Augustus de Morgan introduced the term 'gnomonic projection'. The geometric richness of this projection resides in the fact that it represents the curves that give us the sphere's shortest paths (the geodesics, also known as orthodromics) which, as we have seen, are the great circles, as straight

THALES OF MILETUS (c. 624–c. 547 BC)

Very little is known about the life and work of this Greek philosopher and mathematician. We only know what has been passed down to us through references in works written by later philosophers and historians such as Aristotle, Heredotus or Diogenes Laertius. He taught Pythagoras, and in ancient times was regarded as the first philosopher and the first of the seven sages of Greece. He is credited with a series of geometric results including two that are known by the name of the Thales' Theorem:

1. An angle inscribed in a semicircle is a right angle.
2. If any two straight lines are cut by several parallel straight lines, the segments determined on one of the straight lines are proportional to the corresponding segments on the other straight line.

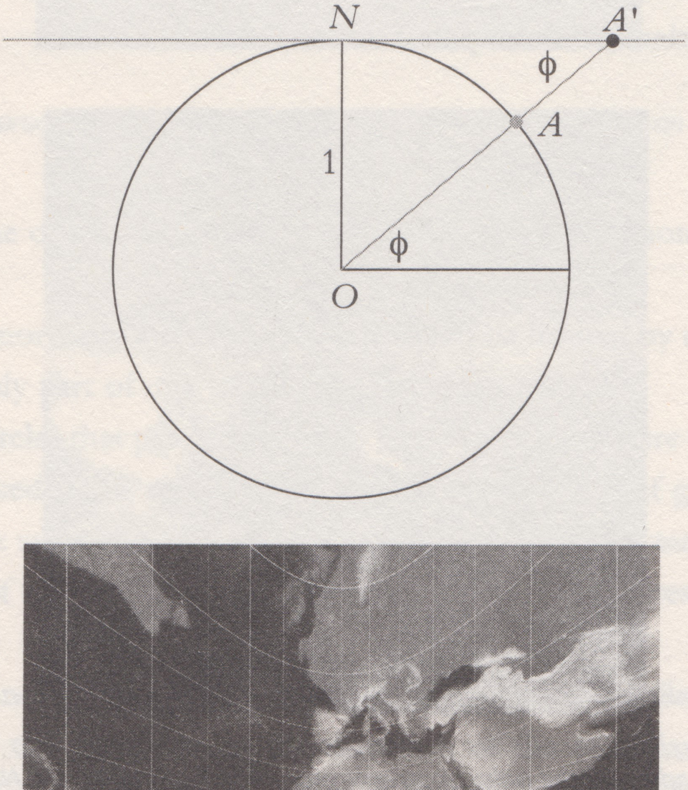


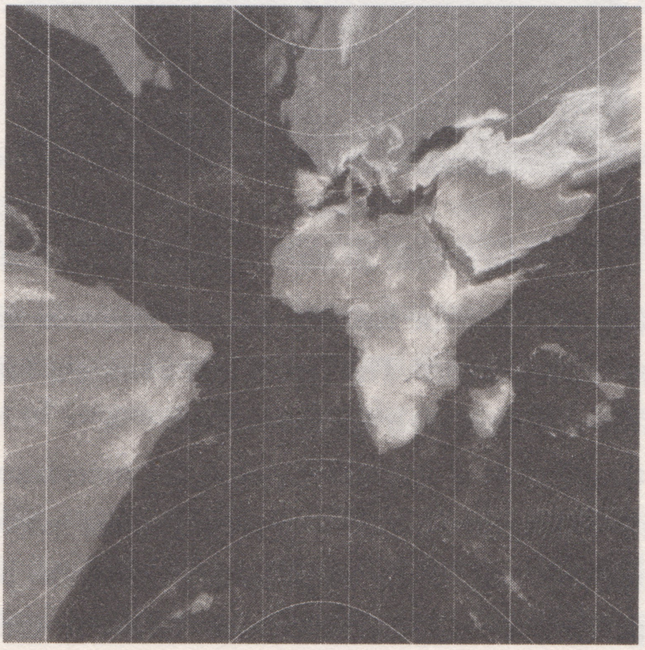
Although we cannot be sure of his true contributions to science, what is certain is that he was the first mathematician to have specific mathematical results attributed to him. He is also recognised as having invented deductive geometry. Various sources describe him as the author of activities of a practical nature, such as measuring the pyramids in Egypt by observing the lengths of the shadows cast by them and by a vertical pole, the prediction of the eclipse of the Sun, and the calculation of the distance of a boat from the shore by means of the proportionality of similar triangles. We owe the story about Thales' wealth to Aristotle. By using his knowledge of astronomy, Thales forecast that the next oil harvest would be plentiful. Consequently, he acquired the oil presses of Miletus and Kios, and by doing so was able, after the harvest some months later, to set the price he wanted for anyone who needed to use them. Thus he managed both to become rich and to put an end to his fellow citizens' rebukes regarding his poverty and the uselessness of philosophy.

lines on the map. Its weak point is the very accentuated deformation noted as we move away from the point of tangency - the central point on the map - which rules it out in cartography as a global map, but not for other purposes, as we shall see later.

Definition and cartographic properties

We shall consider a sphere and a plane tangential to it. The image of a point *A* on the sphere made by means of the central projection is point *A'* on the plane obtained by intersecting it with the straight line that passes through *A* and the centre of the sphere.

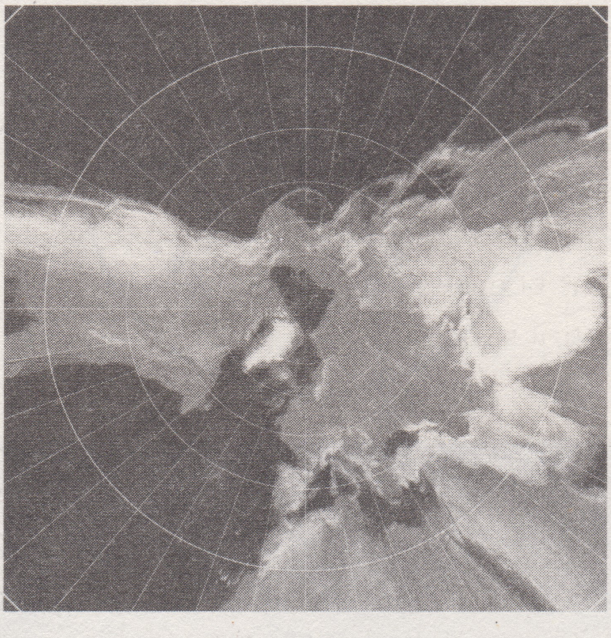




An image of the central or gnomonic projection and map made by using the projection in its equatorial version.

This projection is clearly a geometric projection. If we again imagine the terrestrial sphere as a semi-transparent plastic ball on which the continents and other geographic features are drawn, we can see the image of this map on the white table on which we placed the plastic ball, when we position a light in the centre of the ball and switch it on.

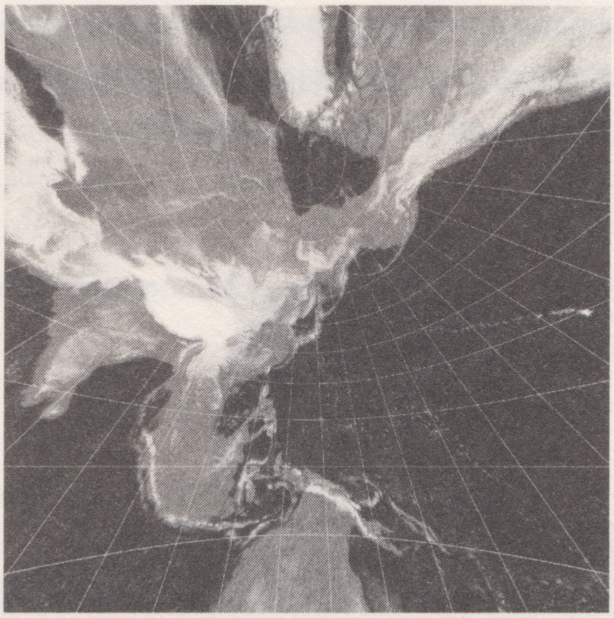
If the point of tangency is one of the poles, then the meridians are radial equally-spaced straight lines emanating from the centre of the map, which represents the pole. The Equator is the infinite, which can never be represented, although, as the map has to have a finite extension, neither shall we be able to represent a whole hemisphere as it would require an infinite plane. The other parallels are represented by concentric circles round the pole.



A map made with the polar gnomonic projection, centred on the North Pole.

If the point of tangency is on the Equator, the meridians are parallel straight lines, though they are not equally spaced. The Equator is a straight line perpendicular to the meridians, and the other parallels are hyperbolae.

If the point of tangency is at any other point, the meridians are radial straight lines that point towards the pole, though they are not evenl spaced. The Equator is a straight line perpendicular only to the meridian on which the point of tangency is situated; the other parallels are ellipses for latitudes nearer to the pole belonging to that hemisphere, a parabola for the latitude of the point of tangency, and hyperbolae for more distant latitudes.

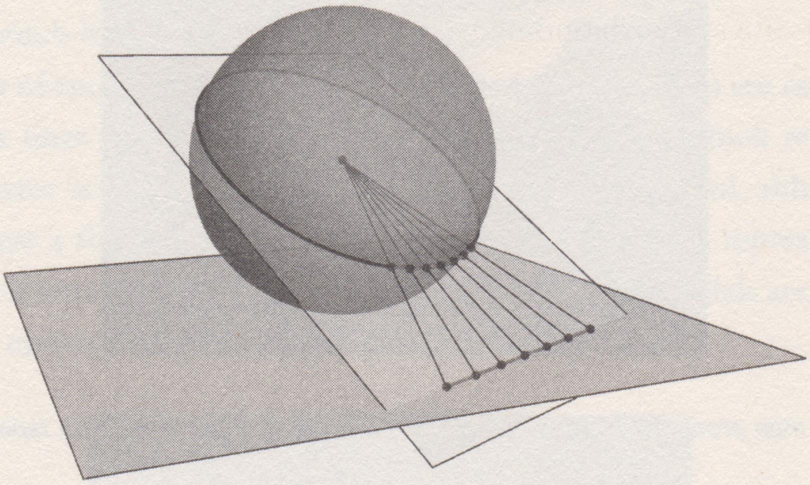


A map produced by the oblique gnomonic projection, centred on Japan.

These are some of the properties of a map created by the gnomonic projection:

1. Its image is normally circular (although it can be limited by other shapes) and it covers only part of one of the hemispheres.
2. The great circles that pass through the point of tangency are transformed into equally spaced radial straight lines (if we take a family of great circles separated by the same angle), while the points that are at the same distance from the point of tangency are transformed into circumferences centred on that point.
3. The shape and distribution of the meridians and the parallels is as described before; and the distortion in the direction of the meridians is $\mu = 1/\sin^2\phi$, while the distortion in the direction of the parallels is $\lambda = 1/\sin\phi$.
4. This projection preserves the geodesics, but not distances, angles or areas.
5. The distortion in areas, shapes and angles, although minor close to the centre (i.e. the point of tangency), is noticeable as we move away from that point.

The proof that the central projection preserves the geodesics is geometric and very simple. The key is to take into account that the geodesics of the sphere are the great circles, that these are produced as the intersection of the sphere with a plane that passes through the centre of it and that, as a consequence, the image of a great circle produced by means of the central projection is the straight line intersection of the plane created by the great circle and the plane of tangency, as can be seen in the image. Thus it is proven that the projection transforms the sphere's geodesics (great circles) into geodesics on the plane (straight lines).



The gnomonic projection preserves the geodesics, and transforms the sphere's great circles into straight lines on the plane.

But, furthermore, it can be proven that this is the only cartographic projection that fulfils this property, as opposed to the case of the areas or the angles. For these properties we can find many projections that conserve them.

With the aim of studying whether the gnomonic projection preserves the areas and/or the angles, we shall look at the distortion produced in meridians and parallels. To do so we'll calculate in essence the Tissot indicatrix of any point on the sphere, that is to say, we take a disc that is sufficiently small (in fact, infinitesimal, so we can consider it on the plane tangential to the sphere at that point) and we shall see what the ellipse transformed by the projection is like.

We start from a sphere of radius 1, which is our terrestrial surface to scale, and the plane of projection *T*, which is tangential to the sphere at the point of tangency (which could be the North Pole, say) and onto which we project part of the hemisphere from its centre. Then, let's consider a point *A* on the sphere of latitude ϕ and a disc *D* of sufficiently small radius *r*, tangential to the sphere at that point.

To see its image, on the plane of projection *T*, produced by central projection, we shall split the study into two phases. In the first, the disc *D* is transformed into a disc *D'*, which lies on a plane parallel to the one of *D* and whose centre is *A'*, the image of A by means of the central projection. By similarity of triangles (Thales' theorem), in the next figure we find that the radius r' of D' fulfils that

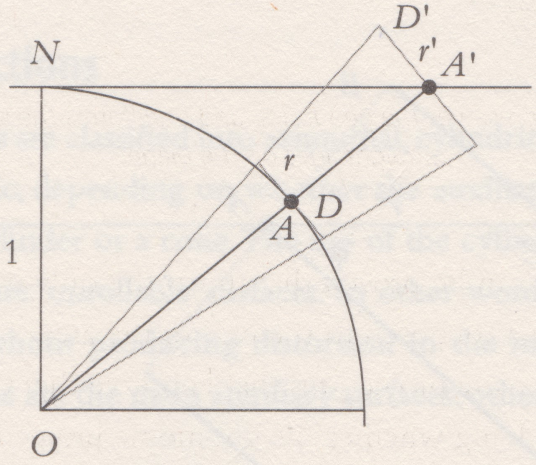
$$\frac{r'}{r} = \frac{|OA'|}{1},$$

and by using elementary trigonometry,

$$\sin\phi = \frac{1}{|OA'|},$$

it can be deduced that:

$$r' = \frac{r}{\sin\phi}.$$



Expansion phase for the gnomonic projection.

Now, the projection of disc D' onto the tangent plane T will give us the required image, which is not a disc now but an ellipse. In the direction 'east–west' the disc D' intersects the plane T, so there is no change on projecting it and, therefore, the semi-axis corresponding to the ellipse will be the radius already calculated:

$$r' = \frac{r}{\sin\phi}.$$

So the distortion on the parallel is:

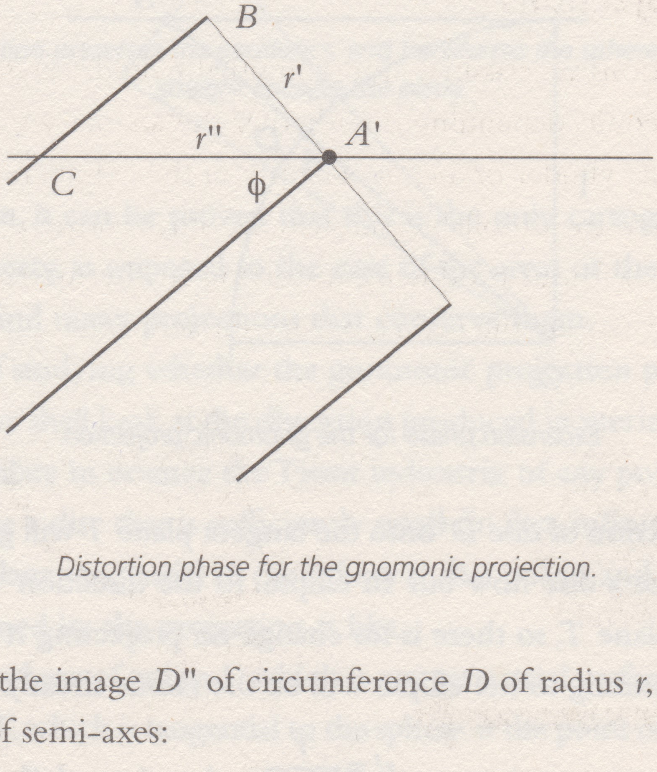
$$\lambda = \frac{1}{\sin\phi}.$$

But let's look at what happens in the direction 'north–south', which will show us the distortion at the meridian. As the radius r' is very small compared with the distance from A' to the origin O, the angle $A'BC$ (next figure) is very close to being a right angle. (Consequently, as r is sufficiently small, we can consider that angle to be a right angle.) Therefore, the segment in the direction 'north–south' of length r' will be transformed into a segment on the plane T of length r'' equal to

$$r'' = \frac{r'}{\sin\phi} = \frac{r}{\sin^2\phi},$$

by again making use of elementary trigonometry. And the distortion on the plane will be:

$$\mu = \frac{1}{\sin^2\phi}.$$



Distortion phase for the gnomonic projection.

Consequently, the image D'' of circumference D of radius r, by central projection, is an ellipse of semi-axes:

$$r' = \frac{r}{\sin\phi} \quad y \quad r'' = \frac{r}{\sin^2\phi}.$$

In conclusion we find that the central projection does not preserve areas since, on account of what was mentioned above, the distortion in meridians

$$\mu = \frac{1}{sin^2 \phi}$$

should have been the inverse of the distortion in parallels

$$\lambda = \frac{1}{\sin \phi},$$

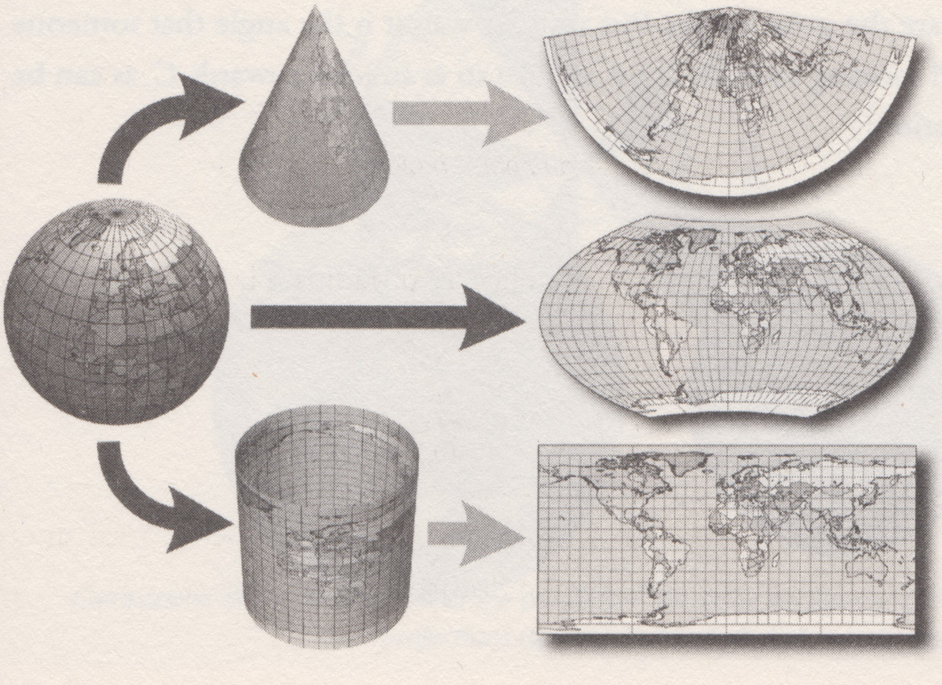
which now is not true:

$$1 \neq \lambda\mu = \frac{1}{\sin^3 \phi}.$$

Neither does it preserve the angles, as the distortion in meridians and parallels is not the same.

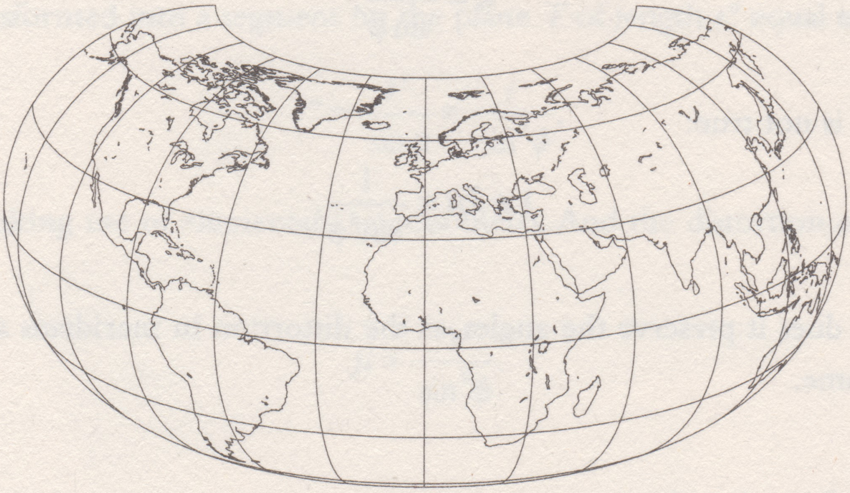
Azimuthal projections

Geometric projections are classified into azimuthal, cylindrical (studied in the previous section), and conic, depending on whether the auxiliary surface used for projection is a plane, a cylinder or a cone. The use of the cylinder and the cone is due to the fact that they are 'unrollable' surfaces. In other words, they can be unrolled into a flat surface without producing distortion in the metric properties in the process. Although these are the main auxiliary surfaces, other intermediate auxiliary



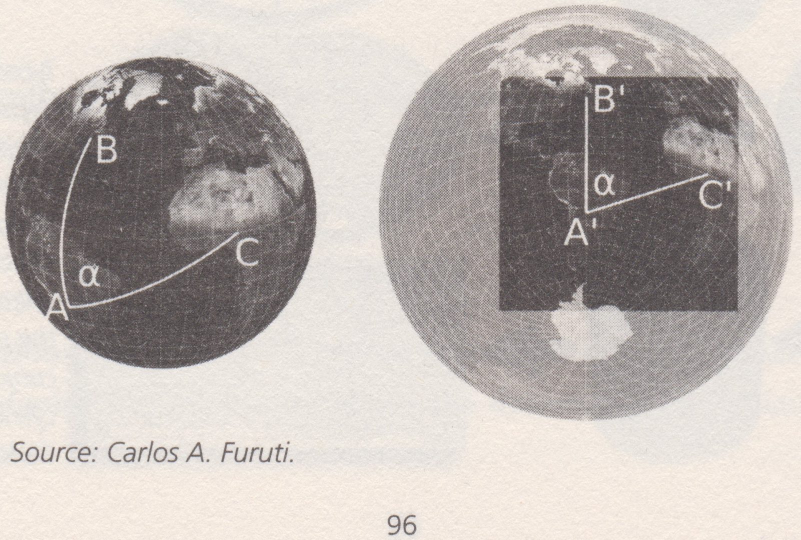
Types of projections: depending on whether the surface on to which the projection is directed is a cone, a cylinder or a plane, the projection is conic, azimuthal, or cylindrical (source: Carlos A. Furuti).

surfaces can be used, and not necessarily ones that can be rolled out (though this is uncommon). For example, the armadillo projection (Erwin J. Raisz, 1943) consists of projecting the sphere onto the surface of a torus (remember that for mathematicians this surface is in the shape of a doughnut or lifebelt) and then projecting it orthogonally, in a certain direction, onto a plane:



Despite the fact that algorithmic projections are expressed by mathematical formulae and have no geometric interpretation, it is normal also to classify them as azimuthal, cylindrical or conic when they share some of their fundamental properties.

Returning to the azimuthal projections, it is surprising that they are not described simply as planar. Where does the term 'azimuthal' come from? Given a point *A* on the terrestrial surface and another two points *B* and *C*, the azimuth of *B* to *C* is the angle formed by the curves of least distance joining *A* to *B* and *A* to *C*, which we already know are the great circles. Put another way, it is the angle that someone situated at point *A* looking towards *B* has to turn so as to look towards *C*, as can be seen in the illustration:



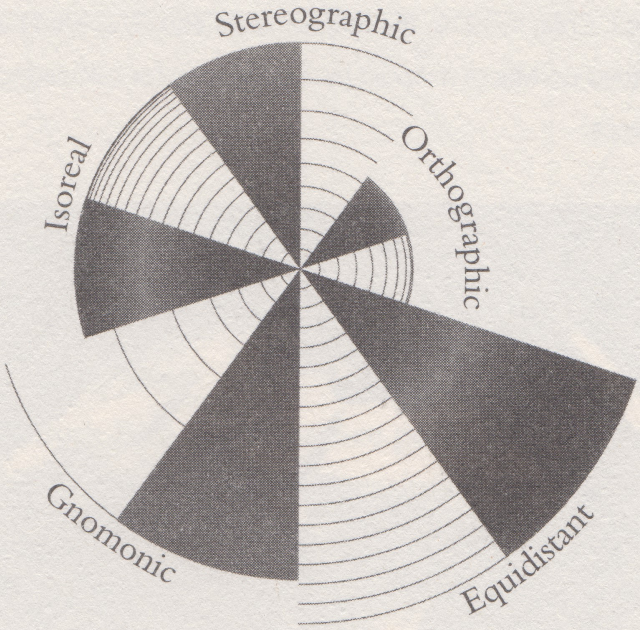
Source: Carlos A. Furuti.

This concept comes from astronomy and nautical studies, and refers to the angle, or length of arc, measured on the distant horizon formed by the cardinal point north, or respectively the south, and the vertical projection of the star on the horizon of the observer.

Therefore, in origin azimuthal projections are those that preserve the azimuth from a fixed point of reference, which is the centre of the map; therefore, they preserve the directions to any other points, but not necessarily distances. Projections that would be called planar were given the name 'azimuthal' on account of the fact that they are produced by projecting directly onto a plane tangential to the terrestrial surface (it could also be a secant plane).

All azimuthal projections in their polar version (with the North or South Pole as their reference point) fulfil the following properties:

1. The meridians are straight lines, and equally spaced (if we consider a grid of meridians separated by a constant angle), that pass through the central point.
2. The parallels are concentric circles centred on the point of tangency; therefore, the different azimuthal projections are determined by the way the circumferences of the parallels are distributed or spaced in the image.



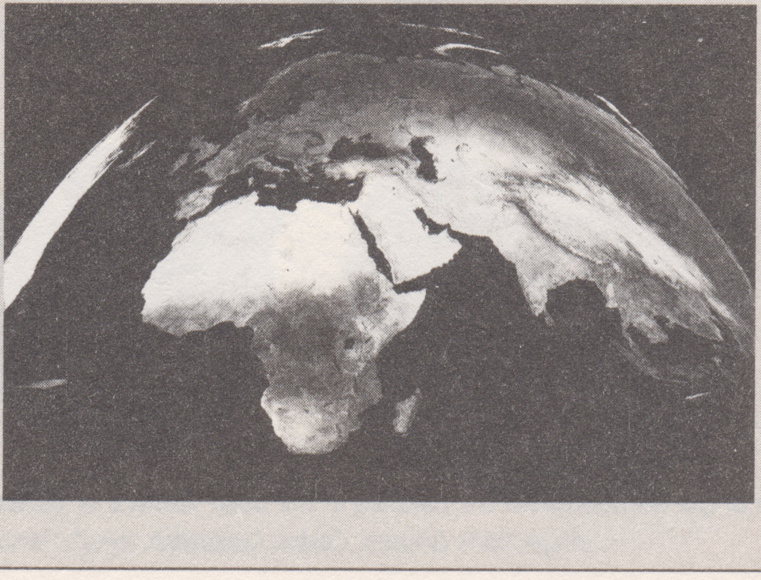
Comparison of the distribution of the parallels in the polar versions of different azimuthal projections (source: Carlos A. Furuti).

In any aspect, not only the polar, the radial straight lines emanating from the centre are the images of the geodesics, i.e. of the great circles, that pass through the point of tangency. But, in general terms, the distances from that point are not preserved (except in the case of the equidistant azimuthal projection). Furthermore, the rest of the geodesics, those that do not pass through the reference point, in general are not preserved except in the particular case of the gnomonic projection, which we are studying here.

MAPS FOR DETERMINING THE DIRECTION OF MECCA

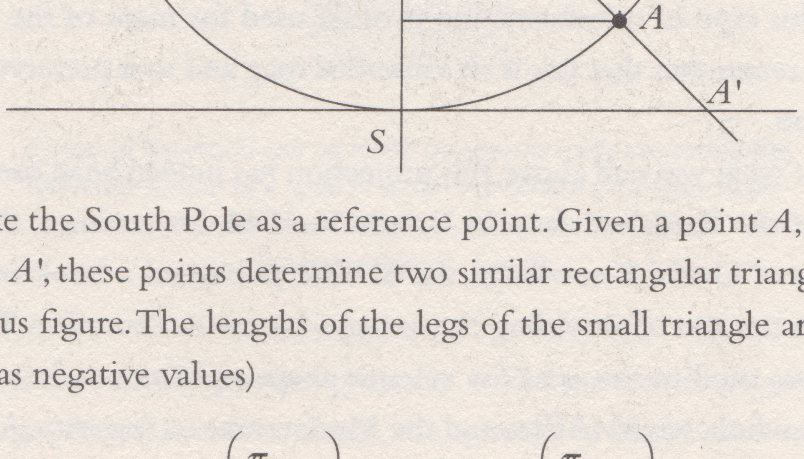
Muslims are required to pray five times a day in the direction of the Kaaba, the cuboid-shaped holy place that symbolises the home of God, located in Mecca. Likewise, mosques have to be correctly oriented. But how can a Muslim finding himself in any part of the world, or who has to build a mosque, work out the direction of Mecca? One possibility is to take a stereographic projection with Mecca in the centre because, as this map is azimuthal and conformal, a straight line can be drawn between the mosque and Mecca and then the angle between it and the meridian can be calculated and thus we can turn this angle with respect to the north to find the direction in which to pray. One of the problems with this map is that its meridians are curved, which makes it a little difficult to calculate the angle.

An alternative idea would be to use retro-azimuthal projections, that is, those from which the direction from any point on the Earth to a fixed point (and not the opposite way around as in the azimuthal) is correct. On the retro-azimuthal projection produced by the British cartographer James I. Craig in 1910, the meridians are equally spaced parallel straight lines. If we used one centred on Mecca it would be a good map for calculating the *qiblah*, that is, the direction of Mecca.



The distortion produced by this family of projections (geodesics, area, angles, shape) is small close to the point of tangency (or around the intersection circle in the case of the secant type) and it gets larger and larger as we move away from there. These projections have the advantage of showing us the image of the Earth as seen from space. Classical geometric projections belonging to this family are the orthographic, gnomonic and stereographic projections, while other more sophisticated projections are the Lambert equal-area azimuthal or equidistant azimuthal. These maps are useful, for instance, in oceanography, long-distance navigation, defence and military issues, and also in tourism, as they are projections focused on a particular spot and preserve the geodesics that pass through it. They produce too much distortion to be used as geographic maps.

Bearing in mind what has been mentioned above, to draw the grid of meridians and parallels for an azimuthal map in its polar version, we have to position the centre (whether it be the North or South Poles), to draw a family of equally spaced straight lines that pass through the centre of the map and which correspond to the meridians, and to draw a family of concentric circumferences corresponding to the parallels. The information we need to know, therefore, is the distance at which we have to draw each of those circumferences. For example, we can calculate the distance in the case of the gnomonic projection.



Let us take the South Pole as a reference point. Given a point A, of latitude ϕ, and its image A', these points determine two similar rectangular triangles, as shown in the previous figure. The lengths of the legs of the small triangle are (remember that here ϕ has negative values)

$$R\cos\left(\frac{\pi}{2}+\phi\right) \quad \text{and} \quad R\sin\left(\frac{\pi}{2}+\phi\right),$$

while for the big triangle they are R and $r(\phi)$, the distance from A' to the centre. Thales' theorem, again, tells us that

$$\frac{R\cos\left(\dfrac{\pi}{2}+\phi\right)}{R\sin\left(\dfrac{\pi}{2}+\phi\right)}=\frac{R}{r(\phi)},$$

from where

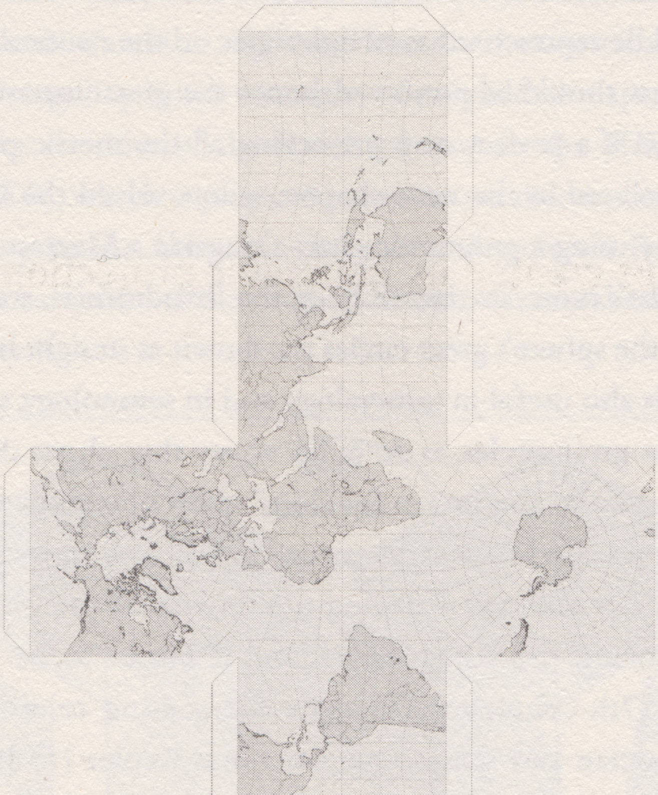
$$r(\phi)=R\tan\left(\frac{\pi}{2}+\phi\right)=-R\cot(\phi).$$

Therefore we are now ready to draw the grid of meridians and parallels of the central projection.

The use of maps created by gnomonic projection

As we said earlier, the central projection is not suitable for designing world maps. A double circle map would be needed, one part for one hemisphere and the other part for the opposite hemisphere. That would not be a problem if it were not that each of the maps would not be able to cover its whole hemisphere and, even worse, the deformation becomes more accentuated the more we move away from the centre of the map and the larger the part of the hemisphere that we want to cover. When we see any gnomonic map the huge distortion in distances, areas and shapes is immediately evident. However, for an angular distance of up to 30° from the point of tangency, a map produced by this projection is quite a good map and exhibits little distortion. This type of transformation is often used for maps of the polar regions. Additionally, remember that this is an azimuthal map and so it preserves the azimuth from its centre.

In spite of what was said above, this projection has indeed been used to represent wide areas of the Earth's surface. In Britain in 1844, for instance, the Society for the Diffusion of Useful Knowledge, the SDUK, published a book of maps in two volumes with the aim of providing cheap maps for use in education. The gnomonic projection was used in the atlas for celestial maps and for a series of six maps to represent the whole world (Africa and the Mediterranean, America, Asia and part of Australia, Oceania and the two poles). These six maps corresponded to the six faces of a cube in which the terrestrial sphere was circumscribed and onto which the projection was effected geometrically from the centre of the sphere, thus producing a series of six maps covering the whole Earth. Another example of a world map in six parts produced in a similar way was published in Weimar in 1803 by the cartographer C.G. Reichard (1758–1837). In 1836 the mathematician Augustus de Morgan published a book with the extensive but explanatory title: *An Explanation of the Gnomonic Projection of the Sphere and of Such Points of Astronomy as are Most Necessary in the Use of Astronomical Maps: Being a Description of the Construction and use of the Larger and Smaller Maps of the Stars; as also of the Six Maps of the Earth.*



A map of the world can be made by gnomonically projecting the reference sphere to a cube in which the sphere is circumscribed, and afterwards unfolding the cube onto the plane (source: Carlos A. Furuti).

In any case, the most important property of the gnomonic projection is that it preserves the geodesics. In other words, the sphere's orthodromics are represented as straight lines on the plane. Therefore, there is an obvious advantage in this type of projection for navigation. Equipped with a navigation chart of the gnomonic type, navigators of a ship or aeroplane who want to calculate the shortest path between two points on the planet, the origin and the destination, simply need to draw the straight line between the images of the two places on the map and they can see which path to take in order to cover the minimum distance separating them. For this reason, gnomonic nautical charts are widely used in navigation. In particular, the United States Hydrographic Office has published this type of chart for all the oceans, and similar agencies in many other countries have also used this projection. In many treatises on navigation, methods for navigating 'by following the great circles' and calculation of distances are explained by making use of gnomonic nautical charts.

The two main conditions in navigation, both maritime and aerial, are that the great circles should be represented as straight lines on the nautical chart and that the angles, the directions, should be preserved; hence the great importance that it would mean for navigation if a perfect map preserving all the metric properties could be designed. As we shall see in the next chapter, sailors solved the lack of an accurate map of the Earth by using a gnomonic chart alongside a Mercator map. The latter is conformal and, what's more, its rhumb lines, the loxodromics, are shown as straight lines. The fact that the sphere's great circles are shown as straight lines means that the central projection is also useful in mineralogy, and in seismology, since seismic waves propagate along the great circles, as do radio waves; thus charts drawn up by means of this projection are also important for ships' radio operators. Central projection celestial maps are even used to observe meteorites, as they move along great circles.

In spite of the fact that the 'horologium' is one of the oldest projections, it was rarely used during the Renaissance, and it began to be used again at the beginning of the 17th century, particularly for making celestial maps. Thus, the German mathematician and astronomer Johannes Kepler (1571–1630) made use of its equatorial aspect to draw a map of the stars in 1606; the Austrian astronomer Christoph Grienberger (1561–1636) used different aspects to create an atlas of constellations in 1616, and Italian mathematician and astronomer Orazio Grassi (1583–1654) did the same to draw up maps of the stars in 1618. Since then it has been one of the most commonly-used projections for designing celestial maps. The reason for this is that those stars that are in the same great circle in the firmament and which appear to be on the same straight line in the sky, are represented on a gnomonic map on the same straight line. This helps in locating them as well as in studying the sky on maps.

The central projection is the transformation that has been used most frequently in creating polyhedral maps and variations of them, though conformal or equal-area projections have also been used for such maps. The idea is very simple: the terrestrial sphere (the reference globe) is circumscribed in a polyhedron and is projected onto it. In the case of the gnomonic projection, in its simple form, it is done from the centre of the sphere. Thus the image of the Earth is produced on the plane sides of the polyhedron and we can then leave it as it is or open it out onto a flat surface.

The most commonly used polyhedrons are platonic solids, whose sides are identical regular polygons (tetrahedron, cube, octahedron, dodecahedron and icosahedron), though other polyhedrons can also be used, such as the truncated octahedron, the cuboctahedron and others. In polyhedral maps the distortion is greater near the vertices and the edges, and lesser in the central areas of the sides, that is, at the point of tangency. Some examples of polyhedral maps, with the gnomonic projection, are the six sides of the Reichard map or the SDUK map, the Cahill butterfly map (1909), which is projected onto an octahedron and later unfolded, or the Dimaxion map by the American designer and architect Richard Buckminster Fuller, which is a variation of the projection onto an icosahedron, which we shall look at in Chapter 9.

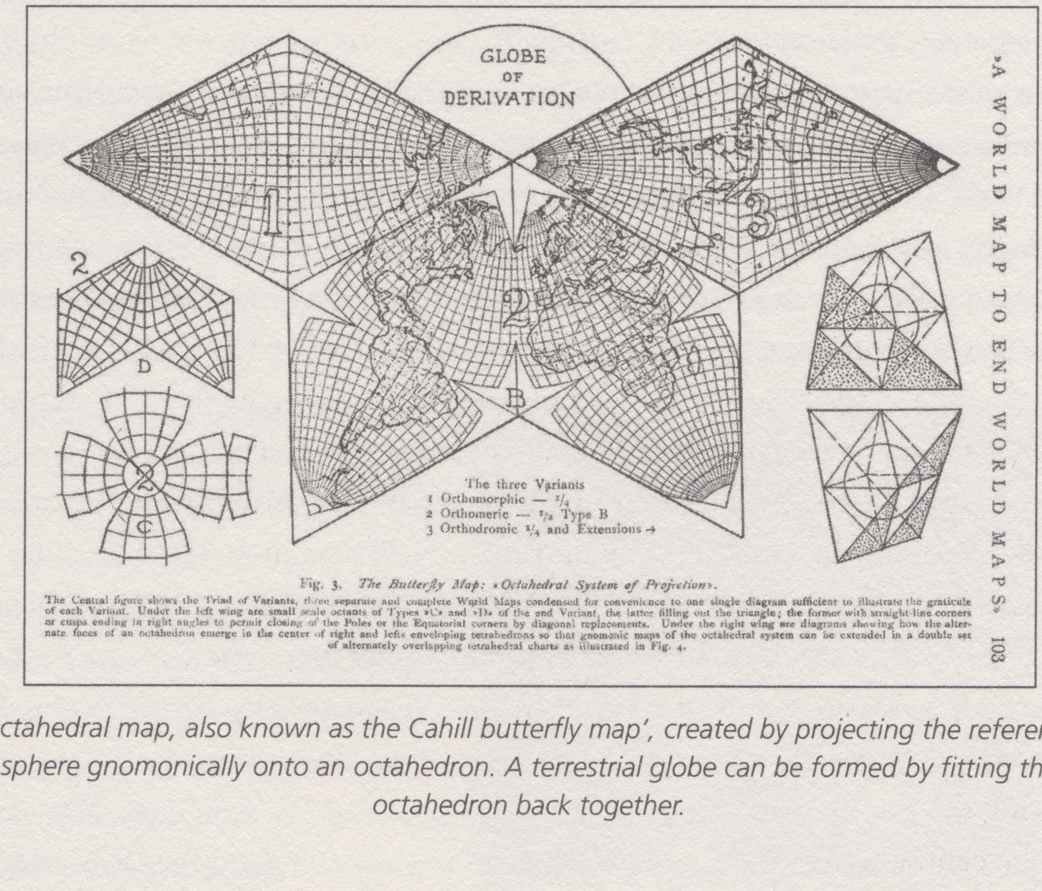

GLOBE
OF
DERIVATION
1
2
3
2
D
2
C
B
The three Variants
1 Orthomorphic — $^{1}/_{4}$
2 Orthomeric — $^{1}/_{2}$ Type B
3 Orthodromic $^{1}/_{4}$ and Extensions →
»A WORLD MAP TO END WORLD MAPS« 103

Fig. 3. *The Butterfly Map: »Octahedral System of Projection«.*

The Central figure shows the Triad of Variants, three separate and complete World Maps condensed for convenience to one single diagram sufficient to illustrate the graticule of each Variant. Under the left wing are small scale octants of Types »C« and »D« of the 2nd Variant, the latter filling out the triangle; the former with straight-line corners or cusps ending in right angles to permit closing of the Poles or the Equatorial corners by diagonal replacements. Under the right wing are diagrams showing how the alternate faces of an octahedron emerge in the center of right and left enveloping tetrahedrons so that gnomonic maps of the octahedral system can be extended in a double set of alternately overlapping tetrahedral charts as illustrated in Fig. 4.

Octahedral map, also known as the Cahill butterfly map', created by projecting the reference sphere gnomonically onto an octahedron. A terrestrial globe can be formed by fitting the octahedron back together.

CELESTIAL MAPS I

In addition to maps of the terrestrial surface there are other maps that have been outstandingly important throughout the history of humankind, and in all cultures: the maps of the sky with its stars and constellations, from the celestial maps of ancient China, India, Mesopotamia or Egypt, which influenced the later Greek, Roman and Islamic maps, to the advanced maps of European celestial cartography. In the 17th and 18th centuries, a great number of celestial atlases were produced in Europe showing the constellations as representations of heroes of Greek mythology, animals (real or imaginary) or diverse instruments. At that time the sky was very important in European society, firstly because it was an essential tool for navigation and secondly because of the popularity of astrology; we have to bear in mind that throughout history many astronomers supplemented their income by also practising as astrologers.

The first atlas of the celestial canopy as a whole, and the one that would prepare the way for the later development of celestial maps, was entitled *Uranometria* (1603) by the Bavarian lawyer and publisher Johann Bayer (1572–1625), which was based on the list of stars drawn up by the Danish astronomer Tycho Brahe (1546–1601). This work appeared prior to the appearance of the telescope, an instrument invented in 1609 by the Italian mathematician and astronomer Galileo Galilei (1564–1642), which contributed greatly to the development of astronomy and stellar maps.

Two of the great achievements of the times of celestial cartography were *Harmonia Macrocosmica* (1660), by German mathematician and cartographer Andreas Cellarius (c. 1596–1665), which was the best known atlas of the 17th century and, for some, the most beautiful of all the celestial atlases ever made, and the *Atlas Coelestis* (1729) by the English astronomer John Flamsteed (1646–1719), who was Britain's first Astronomer Royal based at the Greenwich Observatory.

An image from the work Harmonia Macrocosmica, *by Andreas Cellarius, published in 1708.*

Chapter 7

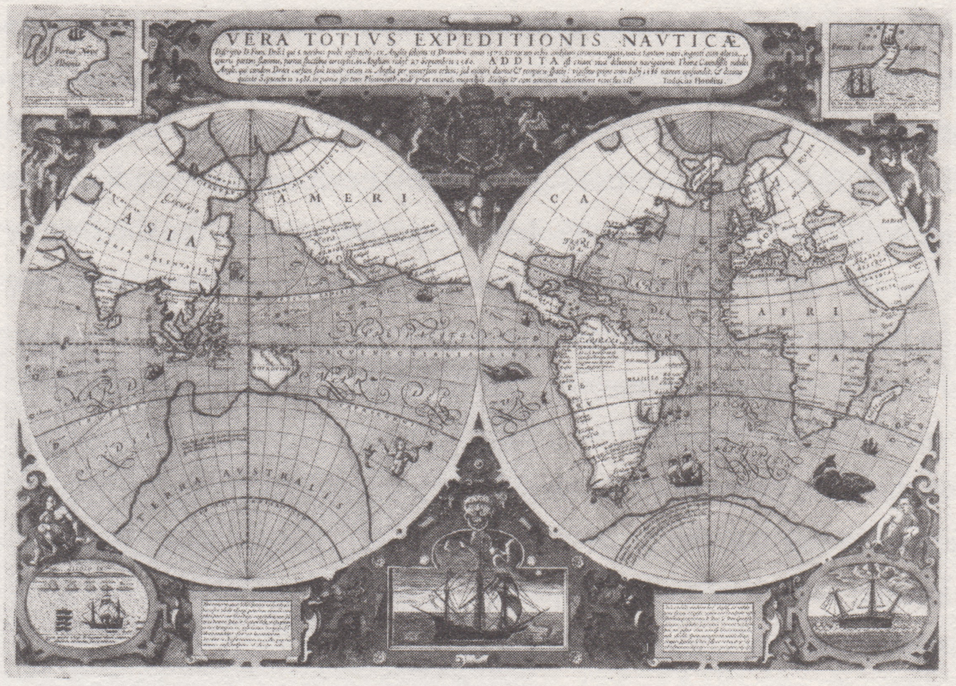
Stereographic Projection

> *The stereographic projection is a graphic method of representing three-dimensional geometric information in two dimensions and of solving problems of three-dimensional geometry. In geology the method is mainly used to solve problems related to the orientation of straight lines and planes, both in crystallography and structural geology. Such problems have more to do with the angular relationship between lines and planes than with their spatial relationships.*
>
> R.G. Park, *Foundation of Structural Geology* (2004)

Stereographic projection has probably been the most common and best known azimuthal projection since it first appeared, and also one of the projections most used in cartography. This projection is commonly attributed to Hipparchus of Nicea, but it is likely to have been known to the Egyptians. The most ancient document that refers to it is a treatise by Ptolemy entitled *Sphaera a planetis projectio in planum*. The original Greek text has been lost, but it has, however, come down to us through the Arabic language. It was the mathematician Maslamah who translated it from Greek into Arabic and this translation has subsequently been used to produce translations and further editions. This piece by Ptolemy appeared in print for the first time as an appendix for the 1507 edition of his work *Geography*, with the title *Planisphaerium Ptolemai*. In this piece, the author describes an instrument for determining the position of the stars on the celestial sphere, the astrolabe, for which the stereographic projection was used. Ptolemy called it the 'planispherium projection', which is the name it held up to the 16th century. (The term *planisphere* continues to be used to refer to celestial charts as this was the projection used to create them.) In the Middle Ages it was also known as the 'astrolabe projection'. The name 'stereographic' is due to the Belgian mathematician François d'Aguillon (1567–1617), who provided a study of the properties of orthographic and stereographic projections in his work *Opticorum libri sex philosophis juxta ac mathematicis utiles* (*Six Books of Optics, Useful for Philosophers and Mathematicians Alike*). The word 'stereographic' comes from the Greek words *stereo*, 'solid body', and *grafia*, 'drawing, painting, representation'. Although the polar version

MASLAMA OF MADRID (c. 950–1007)

Âbû-l-Qâsim Maslama ibn Âhmad al-Faradi al-Hasib al-Qurtubî al-Maŷrîtî was born in Madrid halfway through the 10th century (his nickname 'al-Maŷrîtî' means 'of Madrid'). As a young man he moved to Cordoba, where he was in contact with the group of scientists who had introduced Greek science to Al-Andalus (Islamic Iberia). Maslama eventually formed a scientific school in Cordoba where, on account of the fame he was acquiring (he earned himself nicknames such as 'the Andalusian Euclid' or 'the Prince of the Andalusian Mathematicians', scholars came from all over Andalusia, and so his teachings spread geographically. Outstanding among Maslama's contributions are the translation into the Arabic language of Ptolemy's *Planispherium*, a work which, like the original, has been lost, but which was used for later translations into Latin and Hebrew which kept the comments that Maslama made on the work. The introduction of techniques for constructing the astrolabe, a topic on which he wrote a small treatise; the adaptation of the tables of Al-Jwarizmi and Al-Battani to the meridian running through Cordoba, making them more exact and easier to handle; a handbook of commercial arithmetic; a work on astronomy, and the determination of the longitude of the star named Pumpkin (today, Regulus).

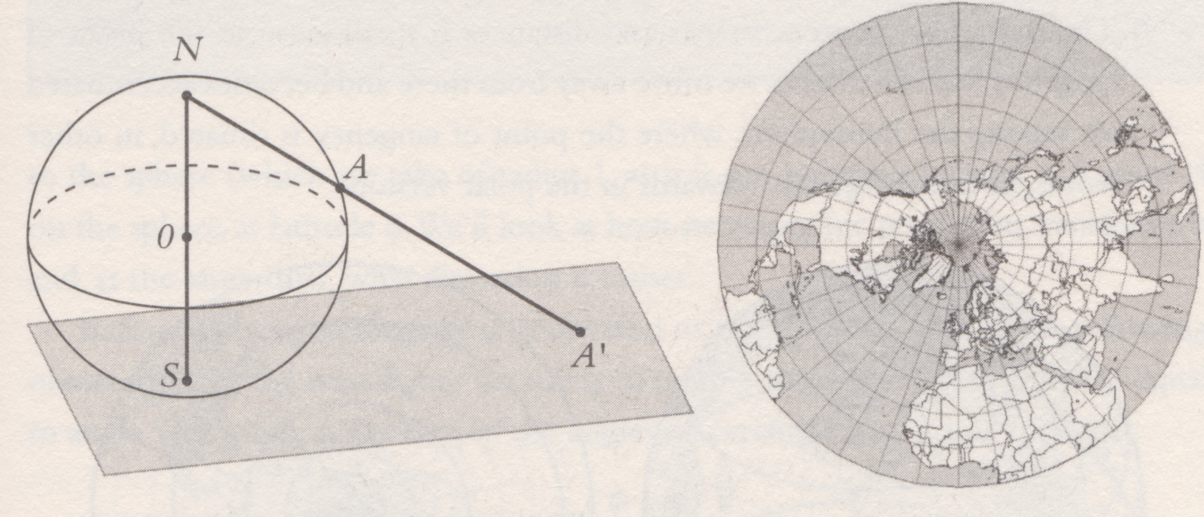


A world map in two hemispheres, Vera totius expeditionis nauticæ *(1595) by Joducus Hondius (1563-1612), produced by stereographic projection. It shows the routes of the first voyages round the world made by the English, both by Sir Francis Drake, between 1577 and 1580, and by Thomas Cavendish, between 1586 and 1588.*

of this projection was known and used in ancient times for making celestial maps, towards the end of the 16th century and during the 17th and 18th centuries it was used regularly to represent the Earth in two hemispheres.

Definition and cartographic properties

The stereographic projection is defined as follows. Let's consider the sphere with a plane tangential to it at a point *S* (for example, the South Pole), and let's carry out the projection from the antipodal point *N* (the North Pole, if the South Pole is the point of tangency). The image of a point *A* of the sphere produced by means of the stereographic projection from *N* is point *A'* on the plane that is obtained by intersecting the plane with the straight line that passes through points *A* and *N*, as the figure shows. Put another way, it is the shadow cast by a light bulb fitted in the north onto a table, or a sheet of paper, if the plastic terrestrial sphere is balanced on the table right at the sphere's South Pole.



On the left, definition of a stereographic projection; on the right, a map drawn up by polar stereographic projection, with the South Pole as the focus of projection.

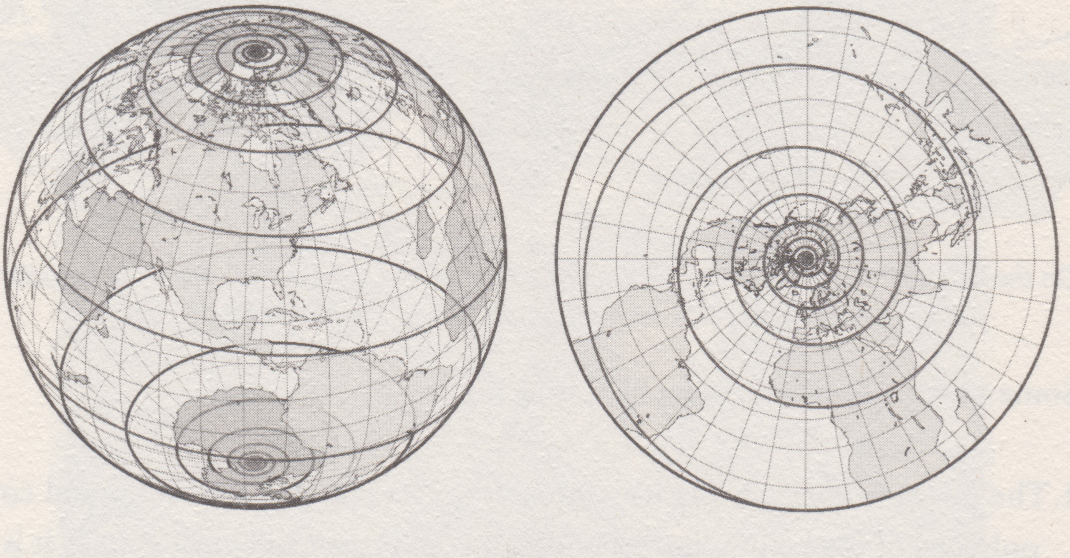
Some properties of stereographic projection are:

1. The image is normally circular, due to the fact that it is azimuthal, and covers only one of the hemispheres. Though it could cover still greater areas, it is not recommended due to the increase in distortion.
2. The distortion in meridians and parallels is

$$\mu = \lambda = \frac{1}{\cos^2\left(\frac{\varphi}{2} + \frac{\pi}{4}\right)};$$

therefore, the projection is conformal, in other words, it preserves the angles, though it does not preserve the geodesics, or the area, or – of course – the distances.

3. As an azimuthal projection, it does preserve the geodesics that pass through the point of tangency, that is to say, in the case that one of the poles is the focus of projection, the meridians are shown as straight lines that pass through the centre of the map.
4. All the meridians and parallels, though more generally all the sphere's circumferences - the great circles and the normal ones - project as circumferences on the plane except the circumferences that pass through the point of tangency, which are transformed into straight lines (this is a particular property of mathematical projections called inversions, and a stereographic projection is an inversion).
5. The loxodromics, or rhumb lines (that is, the curves on the sphere that form a constant angle with the meridians) are transformed into logarithmic spirals.
6. The distortion in areas, shapes and distances is small close to the point of tangency but increases as we move away from there and becomes accentuated on leaving the hemisphere where the point of tangency is situated, in other words, from the Equator onwards in the polar versions.

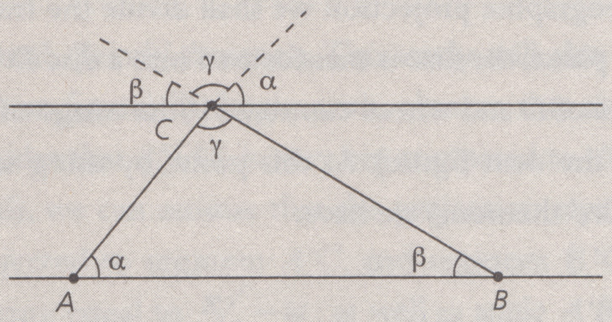


The loxodrome, or rhumb line, seen on the terrestrial sphere and on a map produced by stereographic projection centred on the North Pole (source: Carlos A. Furuti).

Next, to calculate the distortion produced by stereographic projection, we'll proceed in a similar fashion to how we did in the case of central projection. Let's consider a disc D of radius r which is sufficiently small (infinitesimal) and tangential

THE SUM OF THE ANGLES OF A TRIANGLE

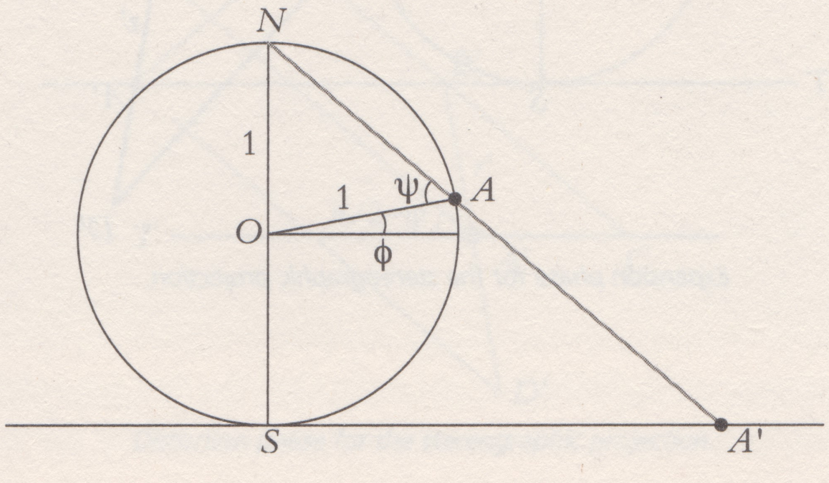
We all know that the sum of the interior angles of any triangle is 180° (or π radians), half of a complete turn. This classical result of Euclidean geometry appeared in the *Elements* (Proposition I–32) by one of the fathers of geometry, the Greek mathematician Euclid of Alexandria (c. 325–365 BC). It is an extremely simple and beautiful proof. Given a triangle *ABC* like the one in the diagram, we draw the straight line parallel to line *AB* which passes through *C*. As they are parallel, with the straight line *AC* they form equal angles (angle α). For the same reason, the angles of the two lines parallel to straight line *BC* are equal (angle β). And as the straight lines *AC* and *BC* are secant, the angle γ and its opposite are equal. Finally, the sum of the three angles of the vertex *C*, which is equal to the sum of the angles α, β and γ of the triangle, is an open angle, of 180°.



to the sphere (which we take of radius 1, as it is our reference sphere) at a point *A* on the sphere, at latitude ϕ. We'll look at how stereographic projection transforms it and, at the same time, what distortion it causes.

Before applying stereographic projection to our disc, let's look at the following observation. In the next figure we use ψ to denote the angle *ONA*, which is equal to angle *OAN* and, as the sum of the angles of a triangle is π, we deduce that:

$$\psi=\frac{\phi}{2}+\frac{\pi}{4}.$$

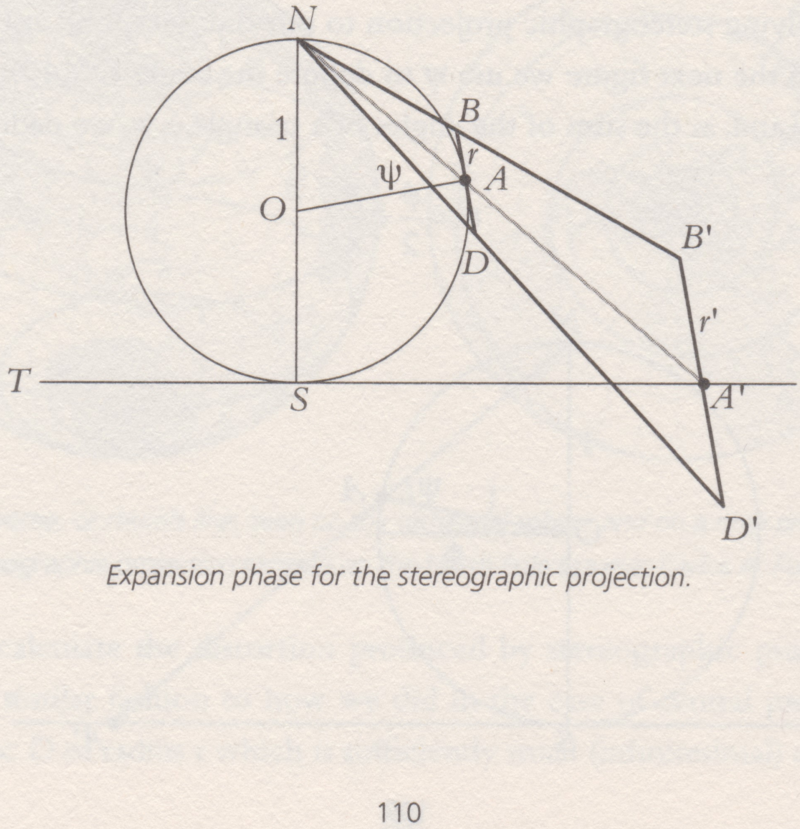


And, in accordance with angle ψ, we know that the distance between N and A is $|NA| = 2\cos\psi$, for which we use the law of cosines used in trigonometry (given a triangle of sides a, b and c and angle α, opposite side a, then $a^2 = b^2 + c^2 - 2bc\cos\alpha$). While by simply taking into account the definition of cosine it is observed that the distance between N and A', the image of A through stereographic projection, is:

$$|NA'| = \frac{2}{\cos\psi}.$$

As before, to get a better understanding of the transformation of the disc through stereographic projection we shall divide the transformation into two phases. In the first phase, the disc is transformed into a disc D' which lies on a plane parallel to the plane of D and whose centre is A', the image of A through stereographic projection (see the next figure). At this point, by using similarity of triangles reasoning (the Thales' theorem) we have:

$$r' = r\frac{|NA'|}{|NA|} = \frac{r}{\cos^2\psi} = \frac{r}{\cos^2\left(\frac{\varphi}{2} + \frac{\pi}{4}\right)}.$$



Expansion phase for the stereographic projection.

The second phase consists of projecting disc D' of radius r' onto the plane of projection T. In the direction 'east/west', disc D' and plane T intersect, so the image segment is the same. This also tells us that the distortion on the parallels is

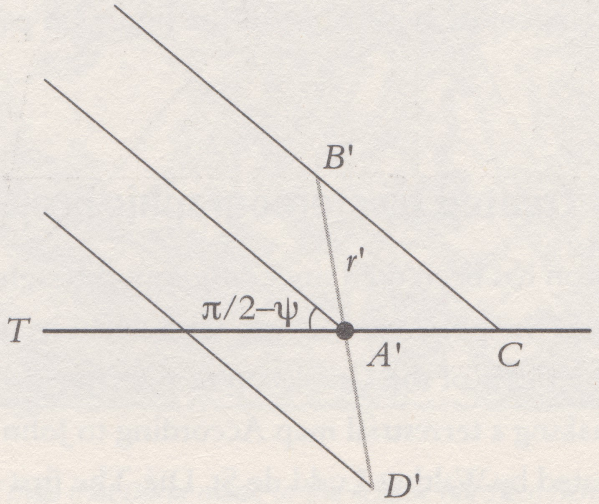
$$\frac{1}{\cos^2\left(\frac{\varphi}{2}+\frac{\pi}{4}\right)},$$

as we have calculated specifically how an infinitesimal segment of length r is transformed on a parallel.

Let's see what happens in the 'North/South' direction. This study will also tell us the distortion on the meridian (next figure). To begin with, observe that angle $SA'N$ is equal to $\pi/_2-\psi$. If we consider that $|NA'|$ is very big compared to r'(we initially considered D as infinitesimal), we can assume that the projection rays are parallel. Then, segment $A'B'$ is projected to segment $A'C$, and segment $B'C$ is parallel to NA', with angle $A'CB'$ being equal to $\pi/_2-\psi$, as well as angle $A'B'C$; consequently, triangle $B'A'C$ is isosceles. Therefore, $|A'C|=|A'B'|=r'$. Therefore we have proved that the image of D after stereographic projection is a disc of radius

$$r'=\frac{r}{\cos^2\left(\frac{\varphi}{2}+\frac{\pi}{4}\right)},$$

so the distortion is the same in any direction and, in particular, in meridians and parallels. That all tells us that stereographic projection is an isogonal projection, in other words, it preserves the angles.



Distortion phase for the stereographic projection.

In 1695, the English mathematician and astronomer Edmond Halley (1656-1742), motivated by his interest in celestial charts and by using the recently introduced tools of calculus, produced and published the first proof that stereographic projection is conformal.

As we pointed out earlier, conformal projections are precisely those that preserve the shapes, but only at a local level and not globally, as can be clearly seen in the Mercator projection or even in the stereographic projection. When we attempt to draw the outline of a country or the course of a river, what determines their shape is the variation of the direction in which we are drawing. Mathematically we would say that it is the variation of the vector tangent to the corresponding curve. For that reason the property of preserving the angles is the property responsible for locally preserving the shapes. An illustrative example could be Greenland, where the difference between its actual shape and the one given on the Mercator map is very big. However, if we look at small areas on the coast, the difference is small.

Designing the grid of meridians and parallels for a map produced by polar stereographic projection comes down to calculating the distance from the centre at which we must situate the circumferences that represent the terrestrial parallels, as in azimuthal projections the meridians are equally spaced straight lines that pass through the centre. Specifically, the radius of the image circumference of a parallel of latitude ϕ is equal to

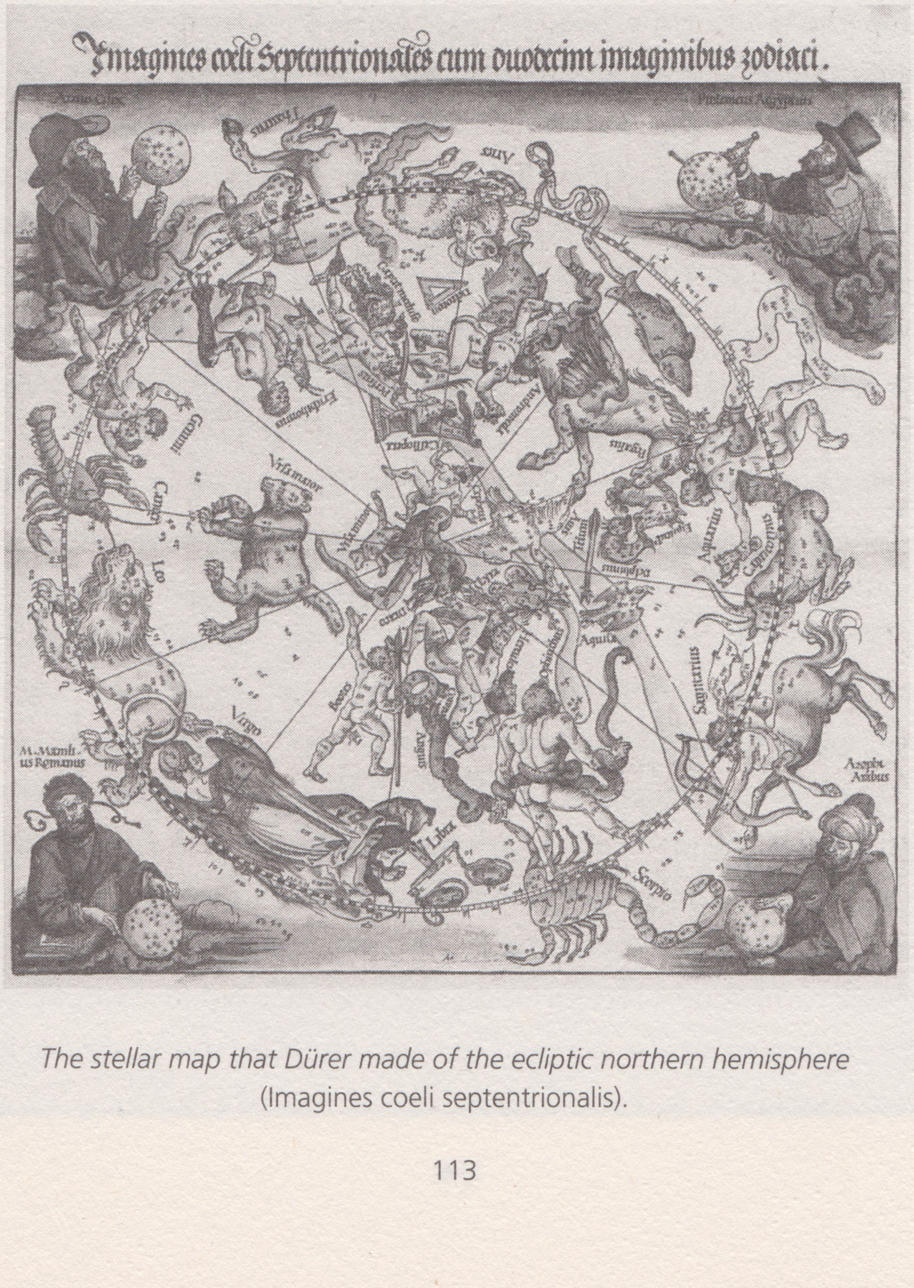
$$r(\varphi) = 2R\tan\left(\frac{\pi}{4} + \frac{\varphi}{2}\right),$$

if R is the radius of the reference sphere. To get that distance we only have to use the definition of the tangent of angle SNA of the figure on page 109 if the sphere has radius R.

The use of maps created by stereographic projection

Stereographic projection has been used since antiquity and right up to our times for designing celestial maps. Its polar version was used exclusively for the purpose of celestial maps from the times of the Greeks up to perhaps 1507, the year in which it was first used for making a terrestrial map. According to John Snyder in *Flattening the Earth*, this was created by Walther Ludd de St. Dié. The first printed stellar maps, using polar stereographic projection, were produced in 1515 by the famous German painter Albrecht Dürer (1471–1528), *Imagines coeli septentrionales cum duodecim imaginibus zodiaci*, for the ecliptic northern hemisphere, and *Imagines coeli meridionales,* for the ecliptic south, which were respectively centred on the ecliptic North and South Poles. Many others followed, such as the two ecliptic celestial hemispheres produced by Halley around 1678, and the *Planisphère contenant les constellations celestes* (published in *Mémoires de l'Académie Royale des Sciences* in 1756), by the French astronomer Nicolas Louis de Lacaille (1713–1762), which showed the stars visible in the equatorial southern hemisphere and would be included in the 1776 edition of the famous *Flamsteed Celestial Atlas*.

Many other stellar maps created by polar stereographic projection were drawn up during the golden age of celestial cartography by great astronomers, cartographers and mathematicians such as Johann Doppelmayer, Pierre-Charles Le Monnier or Jean-Dominique Cassini, and many others.


Imagines coeli Septentrionales cum duodecim imaginibus zodiaci.
Aratus Cilix
Ptolemeus Aegyptius
M. Manlius Romanus
Azophi Arabus
Gemini
Cancer
Leo
Ursa maior
Virgo
Libra
Scorpio
Sagittarius
Aquarius
Capricornus
Aquila

The stellar map that Dürer made of the ecliptic northern hemisphere
(Imagines coeli septentrionalis).

The use of polar stereographic projection for terrestrial maps spread from the beginning of the 16th century onwards. The German humanist Gregor Reisch (c. 1470-1525) used it in his encyclopedia *Margarita philosophica nova* (1512) for a simple map centred on the North Pole. The German cartographer Peter Apian (1495-1552) drew up a small circular map of the northern hemisphere and down to 25° of the southern hemisphere in his *Cosmographicus liber* (1524). In his *Atlas nouveau, contenant toutes les parties du Monde* (1730) French cartographer Guillaume de L'Isle (1675-1726) included two maps of the northern and southern hemispheres created by polar stereographic projection. His nephew, the cartographer Philippe Buache (1700-1773), would produce similar works, in particular his famous *Carte des Terres Australes comprises entre le Tropique du Capricorne et le Pôle Antarctique où se voyent les nouvelles découvertes*

CELESTIAL MAPS II

The stars are spread throughout space, and when we see them at night they look as if they were fixed on a great sphere stretching around us. This is due to the fact that they are at such a great distance from us. This perfect sphere on which the stars appear to be fixed is concentric to the Earth but of an indeterminate radius and is known as the celestial sphere. There are two systems of stellar coordinates, which are defined in a way similar to the terrestrial system (with a celestial Equator, North and South Poles, meridians and parallels, and angles with which we can determine the position of any star in the firmament, analogous to latitude and longitude). One of the systems, which was used more in ancient times, is the ecliptic method. This uses as its reference the plane of the ecliptic – the path of the Sun 'around' the Earth – which is used as the 'ecliptic equator'. The other system, which is more modern, is based on the Earth's Equator. Additionally, the zenith is the highest point of the celestial globe above the observer, and the nadir is the point diametrically opposite.

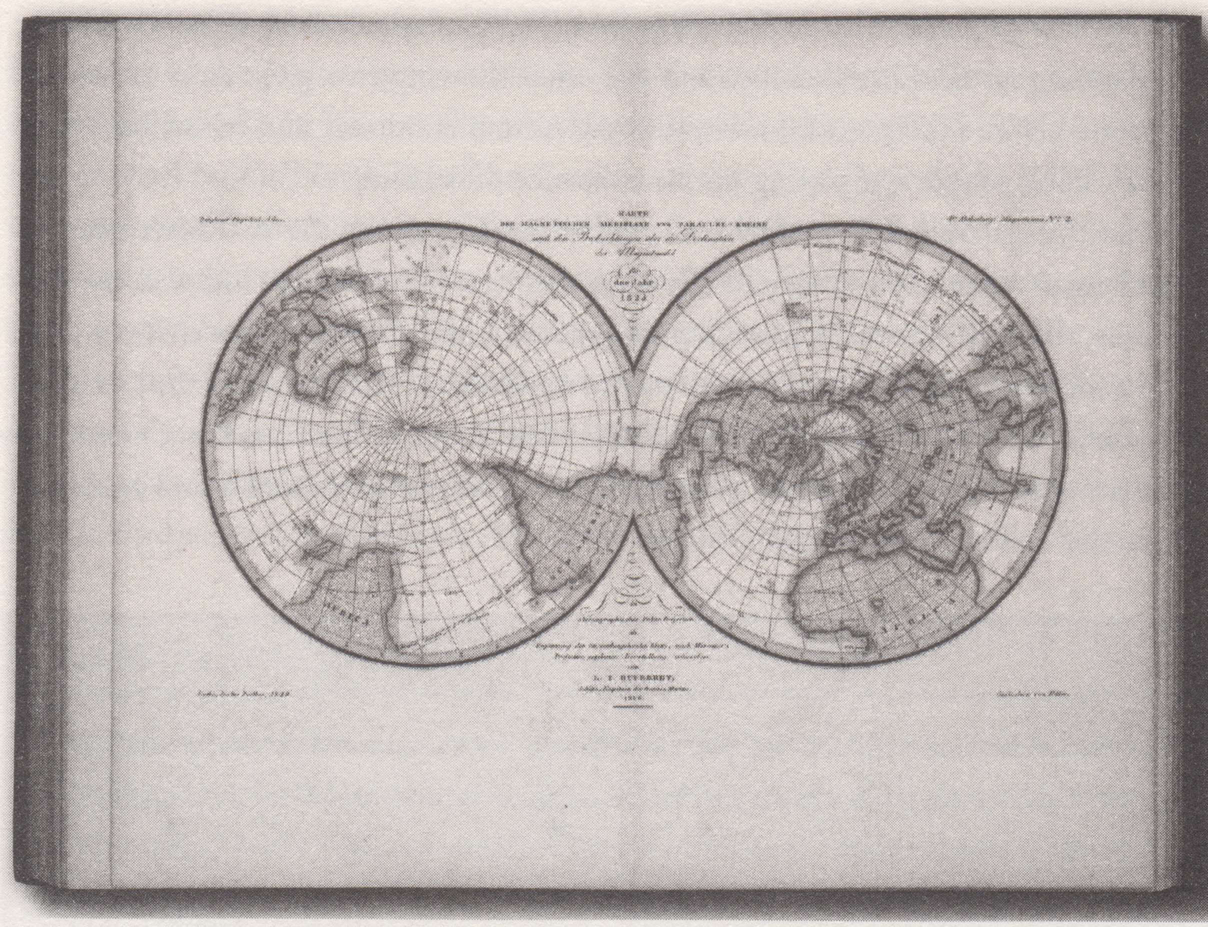
With the positions of the stars and constellations determined by these coordinates, we can project the celestial sphere onto a flat surface and thus obtain a map of the starry sky in a similar way to producing a flat map of the surface of the Earth. The same projections used in terrestrial cartography can be used for celestial projections, and that is in fact what happens. However, the projections that are most used in maps of the stars are the gnomonic, the stereographic and the equidistant, which produce maps which we can centre on the zenith, as the distortion produced is uniform around that central point. What's more, they are azimuthal, which simplifies the problem of finding the direction in which to look from our central observation point in order to see a certain star. Afterwards, depending on our particular interest, *faites en 1739 au Sud du Cap de Bonne Esperance* (1739), of which there are two versions, one showing no land in Antarctica and the other showing two great areas of land, a long time before much was known about the Austral continent, and which has been used by some people speculating on the existence of Atlantis.

These maps were followed by many other similar ones of the two hemispheres. World maps were even designed showing the two hemispheres together as, for example, the map known as *Karte Der Magnetischen Meridiane und Parallel-Kreise* and other similar ones that appeared in the work *Physikalischer Atlas* (1848) by the German geographer Heinrich Berghaus (1797-1884), which was to illustrate the *Cosmos*, or *A Sketch of a Physical Description of the World*, by the German naturalist and explorer Alexander von Humboldt (1769-1859).

we can choose one or other of the three projections – for the stars on the same great circle to be shown on the same straight line, for the angles to be preserved or for the distortions not to be very pronounced.

PLANISPHERE des Étoiles Australes
Dressé par M. l'Abbé DE LA CAILLE
Capricorne

A celestial map of the south polar region by Nicolas Louis de Lacaille, which appeared in the Fortin edition of the Flamsteed Celestial Atlas.

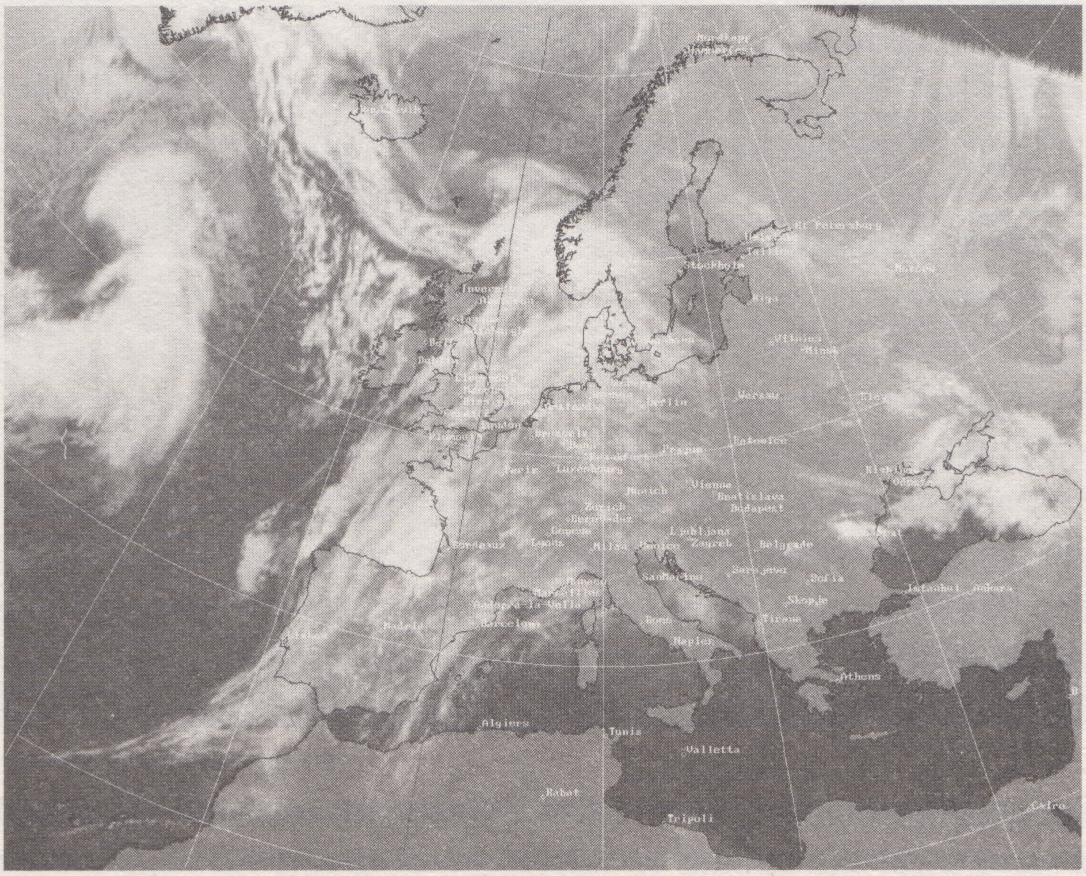


Karte Der Magnetischen Meridiane und Parallel-Kreise, *a map in two hemispheres with stereographic projection by the German geographer Heinrich Berghaus.*

In general, polar stereographic projection has become quite a common projection for drawing up maps of the polar regions because, as areas away from the centre of projection – the corresponding pole – are not the ones being studied, the general metric distortion is very small. In particular, the USGS (United States Geological Survey) and many other international agencies often use it for maps of the Antarctic, with even more reason to do so if the map's theme requires a conformal map, for example, meteorological maps, maps of the Antarctic winds and so on. And likewise for maps of the Arctic Circle. Some specific examples are the National Snow and Ice Data Center or the National Oceanic and Atmospheric Administration (NOAA) in the USA, which also use stereographic projection for the polar regions.

It is also the basis for the UPS (Universal Polar Stereographics) system of coordinates, which, together with the UTM (Universal Transverse Mercator) system, make up a system of coordinates, or projections, for the entire planet. These coordinates define a series of specific zones of an intermediate size for which the chosen projection is fit for purpose, as the distortion is slight. This system of coordinates has thus become the system most used by all the world's agencies for drawing up maps of a certain size. We shall return to this theme in Chapter 9, which is dedicated to the Mercator projection.

The polar stereographic projection is also used for regions of intermediate size that are not polar but which are near to them. Thus they are often used for maps of Russia, of all Europe, or for some European countries such as Sweden, and also for Australia or North America.



A map of the weather in Europe created by polar stereographic projection.

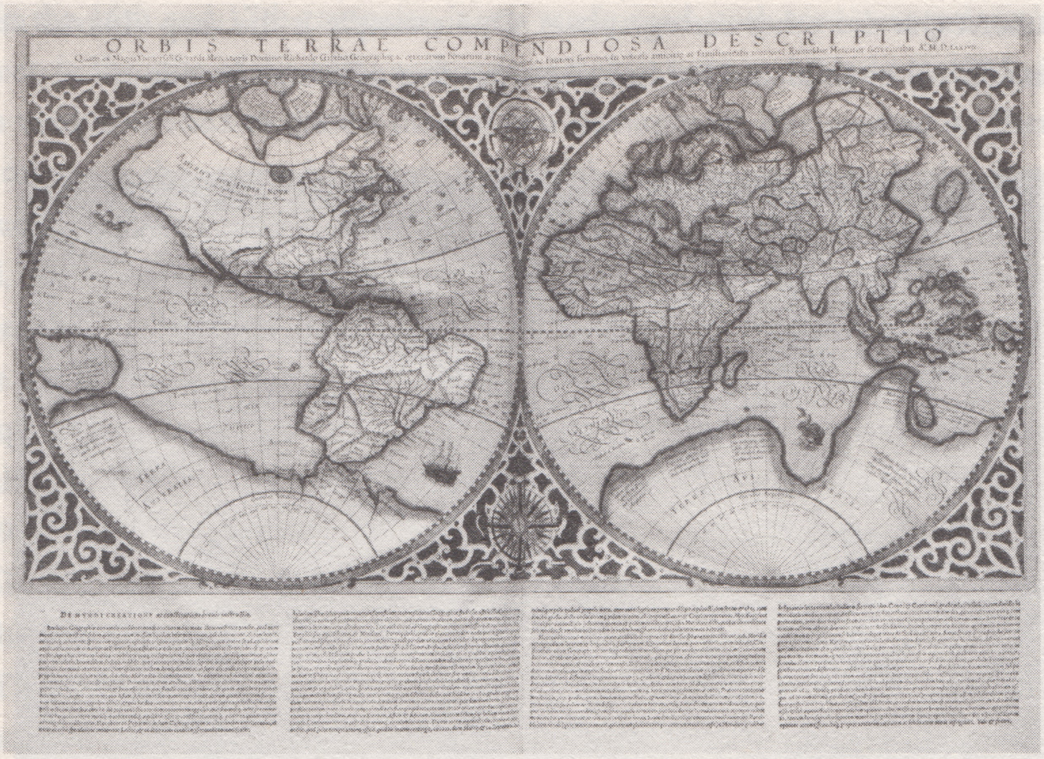
Another aspect of stereographic projection, the equatorial one – that is, the aspect in which the point of tangency with the plane of projection, which is to be the centre of the map, is situated on the equatorial line – was used by the Moorish scholar known as Arzachel when designing the astrolabe. For making world maps it was not used until the 16th century; the simple map in two hemispheres by the French cartographer Jean Rotz was the first of those that still exist.

ARZACHEL (c. 1029–1100)

Abu Ishäq Ibrahim Ibn Yahyà al-Naqqash al-Zarqali was born in Toledo, Spain. His nickname 'Arzachel' (the Latinised version of al-Zarqali) is a reference to his blue eyes (*zarco* = light blue). According to some sources he worked as a metal worker in his father's workshop and received no academic education (some sources even describe him as being illiterate). He began to construct scientific instruments, such as the astrolabe, for the Taifa Kingdom of Toledo, which resulted in his coming into contact with astronomy, a science he then taught himself. After the conquest of Toledo by Alfonso VI in 1085, he moved to Cordoba, where he died in 1100. His contributions to science were in the field of astronomy and of three types: instruments, tables and theoretical studies. Of the instruments he designed and built, the one that brought him most fame was the saphaea, a type of 'universal' astrolabe that could be used at any latitude and continued to be used in Europe until the 16th century. He also built a clepsydra, a water clock which shows the time both during the day and at night, on the banks of the River Tagus. He wrote several treatises for constructing and using instruments, such as the saphaea mentioned above, the Equatorium and the armillary sphere among others. Arzachel adapted the astronomical tables of Al-Jwarizmi and Al-Battani to the coordinates of Toledo, thus creating the famous 'Toledan Tables'. He also collected together his own observations and adapted Greek astronomical information in his Almanac. This document made it possible to fix the position of a planet, the Sun, the Moon and other celestial bodies directly, without the need to make long calculations, and it enabled both solar and lunar eclipses to be predicted. He wrote different treatises in which he gathered together the studies he had made on the Sun, the Moon and Mercury.

But it was Rumold Mercator, the son of Gerardus Mercator and heir to his father's tradition who, with his map *Orbis terrae compendiosa descriptio* (1587) – included by the latter in the 1595 edition of the Mercator atlas – would make the equatorial aspect of the stereographic projection the standard projection for producing world maps in two hemispheres; this would dominate cartography in the following centuries.

Depending on whereabouts on the equatorial line the centre of one of the hemispheres, the point of tangency, is placed, as the other is the antipodal point of the terrestrial sphere, it produces different representations of the world. On maps such as Rumold's, in the *Nova et accuratissima terrarum orbis tabula* (1664) by the Dutch cartographer Joan Blaeu (1596-1673), the son of cartographer Willem Blaeu, or in the *Orbis terrarum nova et accuratissima tabula* (1666), by Peter Goos (which

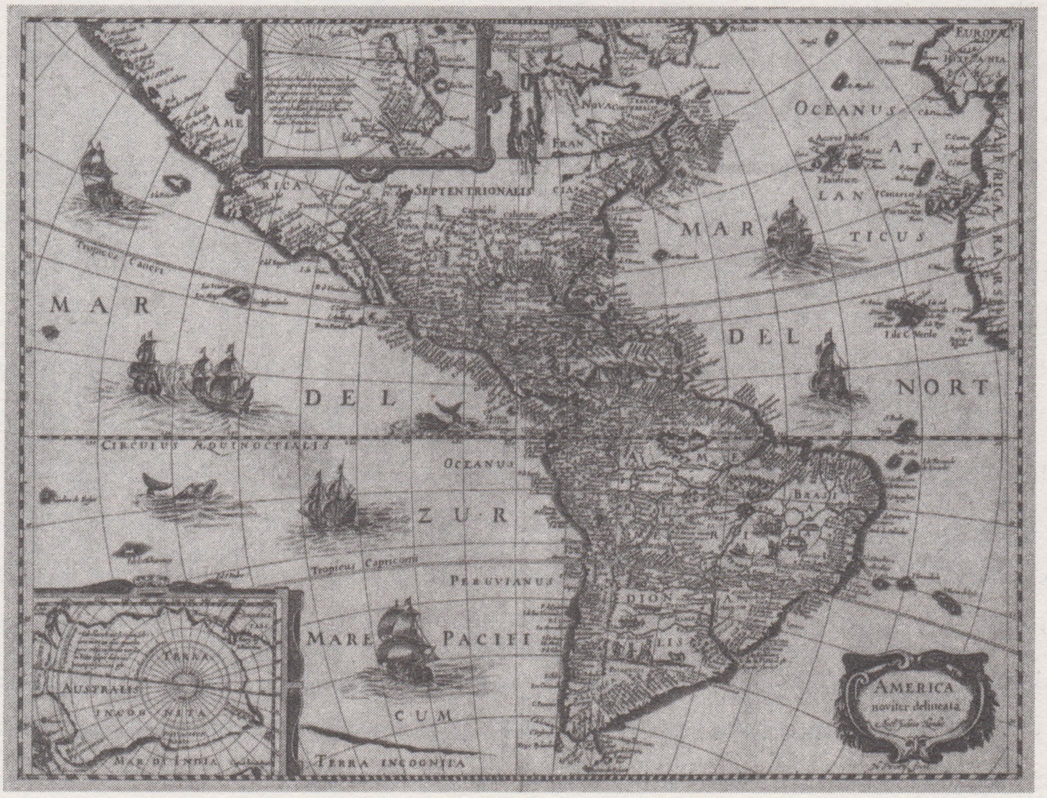


Orbis terrae compendiosa descriptio *(1587), a world map in two hemispheres, produced by an equatorial stereographic projection, by Rumold Mercator. It shows both Australia and Antarctica as cartographers believed them to be shaped at that time.*

show California as if it were an island, a cartographic error that was repeated on many maps in parts of the 17th and 18th centuries), the centre is situated on the meridian that passes through the Caspian Sea, meaning that one of the hemispheres contains Europe, Africa, Asia and part of Oceania, while the other includes the Pacific Ocean and America.

Others, such as Joducus Hondius, who in his *Vera totius expeditionis nauticæ* (page 106) shows the routes around the world taken by Englishmen Francis Drake and Thomas Cavendish, rotated the world 90° and centred the maps on the Atlantic and Pacific oceans to show each of the hemispheres. There were those who situated Europe and Africa in the centre of one of the hemispheres, as did the German Philip Eckebrecht (1594–1667) in a world map he drew up in 1630 for a work on astronomy by the German mathematician and astronomer Johannes Kepler (1571-1630).

The equatorial stereographic projection has also been used since then, and is even still used today, for partial maps of the world. For instance, the German cartographer Johann Baptist Homann (1664–1724) used it not only for the world map in two hemispheres that was typical of those times but also for maps of Europe, Asia, Africa and America, as did other cartographers. One example is the map *America noviter delineata* (1640) by Joducus Hondius, also a product of this projection:


OCEANUS
AT
LAN
TICUS
MAR
DEL
NORT
MAR
DEL
ZUR
MARE
PACIFI
CUM
CIRCULUS AQUINOCTIALIS
Tropicus Cancri
Tropicus Capricorni
SEPTENTRIONALIS
PERUVIANUS
TERRA INCOGNITA
TERRA AUSTRALIS INCOGNITA
EUROPA
AMERICA noviter delineata

The oblique version was first used in the 4th century for maps of the sky by the Greek mathematician and astronomer Theon of Alexandria, who was possibly the last director of the Library of Alexandria and the father of Hypatia. In modern-day treatises on celestial maps, the stereographic projection is one of those recommended, alongside the other azimuthal, gnomonic and equidistant projections, for designing maps of the starry firmament.

The stereographic projection, in its oblique aspect, was proposed for terrestrial maps as an original projection by the Austrian cartographer Johannes Stabius (1450–1522) and promoted as one of his new projections by the German mathematician Johannes Werner (1468–1522) in his translation into Latin of Ptolemy's *Geography*. We have to bear in mind that one of the fundamental problems when suggestions were made for making maps was not to define them geometrically but to find methods to construct them, as there was no availability in those times of the computer programs which we have today. Thus, John Snyder's book *Flattening the Earth* shows some of the construction methods for stereographic projection and other cartographic transformations.

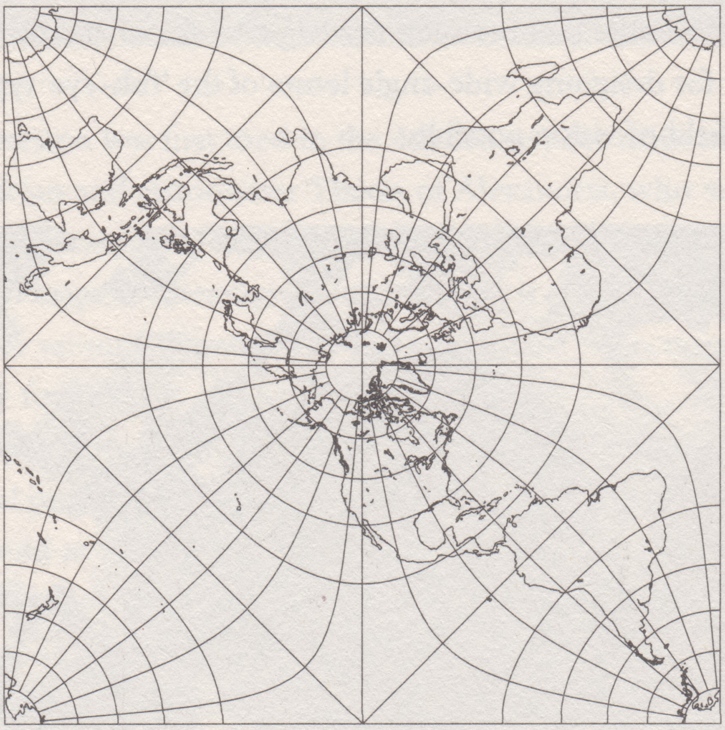
In the 16th and 17th centuries, it was rarely used, with some exceptions, as can be seen in the atlas of the sailor and cosmographist Jacques de Vaulx (c. 1555–1597), who used it in 1583 to draw a map of two hemispheres, one centred on Paris and the other on the antipodal point, and which covers Antarctica. It also appeared in maps of Europe and Asia by the English historian John Speed (1552–1629). And, though it has never been a star player in cartography, oblique stereographic projections are still used today. If we take a look at *Map Projections for Europe* (2003) from the EU's Institute for Environment and Sustainability, for instance, we can find oblique stereographic projection used for maps of Holland, Poland and Romania. Furthermore, it was the basis of the new projection developed by Henri Roussilhe in 1922 and was used both by the USSR and by the USGS.

Getting back to the skies, stereographic projection is also being used, in particular, by the USGS to produce maps of the Moon, Mars and Mercury.

The rich geometry of stereographic projection has made it a useful tool in many branches of science: within mathematics, in areas such as complex analysis, non-Euclidean geometry, differential geometry, analytical geometry or topology, as well as in physics, structural geology or engineering. In crystallography it is used to study the symmetry of crystals because the property of being conformal ensures that the angles between the sides and between the edges of the crystal are maintained, and in photography for designing wide-angle lenses of the 'fish-eye' type made to cover the maximum field of vision possible.

Photos taken with stereographic projection, using a wide-angle fish-eye type lens, have become fashionable in photographic art (source: Alexandre Duret-Lutz).

The conformal applications are particularly suitable in cases in which the angles or direction (or bearings) are an important property, as, for example, in maritime or aerial navigation. It was mentioned above that orthodromes were important for navigation, and, similarly, another outstanding family of curves is the loxodromes (those which form a constant angles with the meridian), as to navigate by them all that is necessary is to maintain the direction shown, for example, by a compass. For that reason, one important conformal projection is the Mercator projection, in which the loxodromes become straight lines and, therefore, are easy to draw on the corresponding map. The preservation of the angles is the reason why they are also used in geodesics, in meteorology (to show, for instance, the direction of the winds, or to show the isobars, which are perpendicular to wind directions) and in oceanography. They are also useful in fields in which circular radiation patterns are used. Let's not forget that conformal applications transform circumferences into either circumferences or straight lines, as in radio transmissions or in seismic waves. For drawing up maps of the brain it is also important for the transformations to be conformal, as can be seen in the studies carried out by the US researcher Monica K. Hurdal (Florida State University), who works in biomedical mathematics.



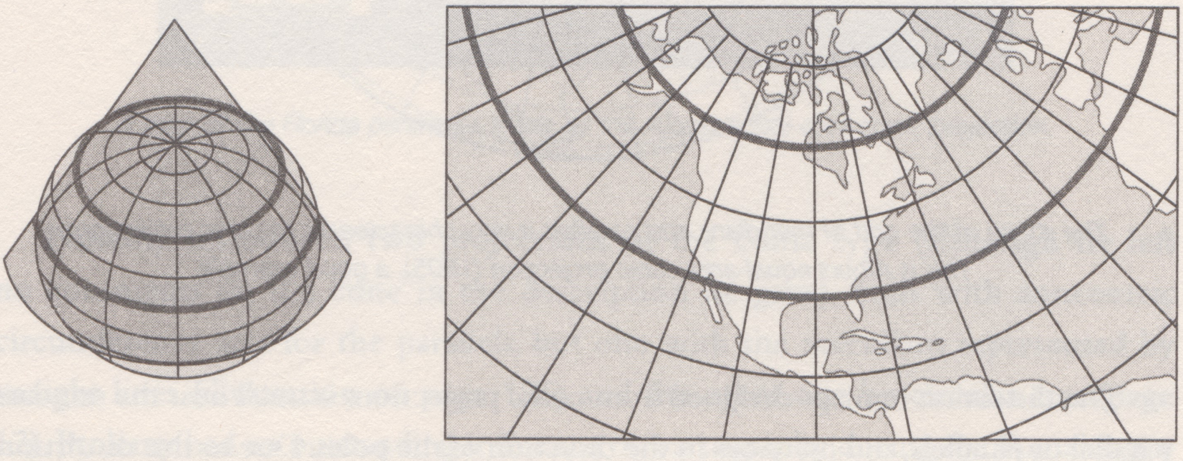
The Pierce quincuncial projection is a conformal transformation, produced by complex analysis techniques from stereographics, which transforms the sphere into a square.

Lastly, as conformal projections preserve shapes at local level, this makes them very suitable for making maps of small regions of the Earth, as will be shown in later examples.

Among the most commonly used conformal projections are the following – the stereographic we have already studied, the Mercator, the Lambert conformal conic or the bipolar oblique conformal conic. Other conformal projections are the Lagrange (presented by Lambert in his 1772 work), those of August and Eisenlohr (introduced around 1870), the square-shaped Pierce quincuncial (1879) and the also square-shaped Goyou (1887), to name just a few.

Conic projections

The most important conformal projection, after the stereographic, which we have just looked at, and the Mercator, which we shall study in Chapter 9, is Lambert's conformal conic projection which, as its name suggests, belongs to the third type of mathematical projections used in cartography – in addition to the azimuthal and the cylindrical ones. Geometric-type conic transformation (and, by extension, the algorithmic type) consists of projecting the reference sphere to a cone tangent or secant to the sphere, and then unfolding the cone into a plane. To do so, the cone is cut along one of its 'meridians' and opened up until it is flat. The reason why a cone is used is that this surface can be opened up into a plane without altering the metric geometry, as was also the case with cylinders. What's more, the circumferences of tangency, or of intersection of the cone with the sphere, are taken as standard lines, or lines of real scale, that is to say, those curves on which the scale is linear.

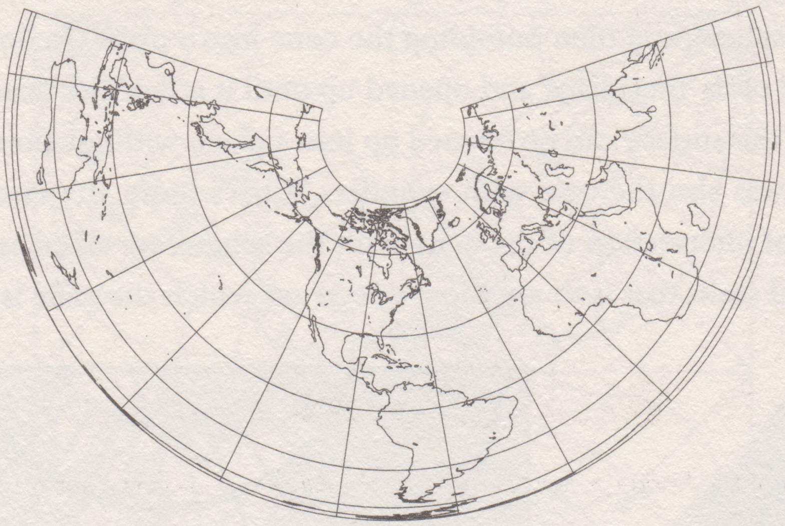


An image projected onto a cone and then unfolded.

All normal versions of conic projections, that is to say, when the vertex of the cone is on the North-South terrestrial axis and the curve of contact with the sphere is a parallel (or two, in the secant case), fulfil the following properties:

1. The meridians transform into straight lines that emanate from a point and are regularly spaced at equal angular intervals. With respect to that of the sphere, the angular distance between the meridians is reduced by a fixed factor, the conic constant.
2. The parallels are projected in arcs of concentric circumferences, which cut the meridians at a right angle. The distortion is constant along each parallel.

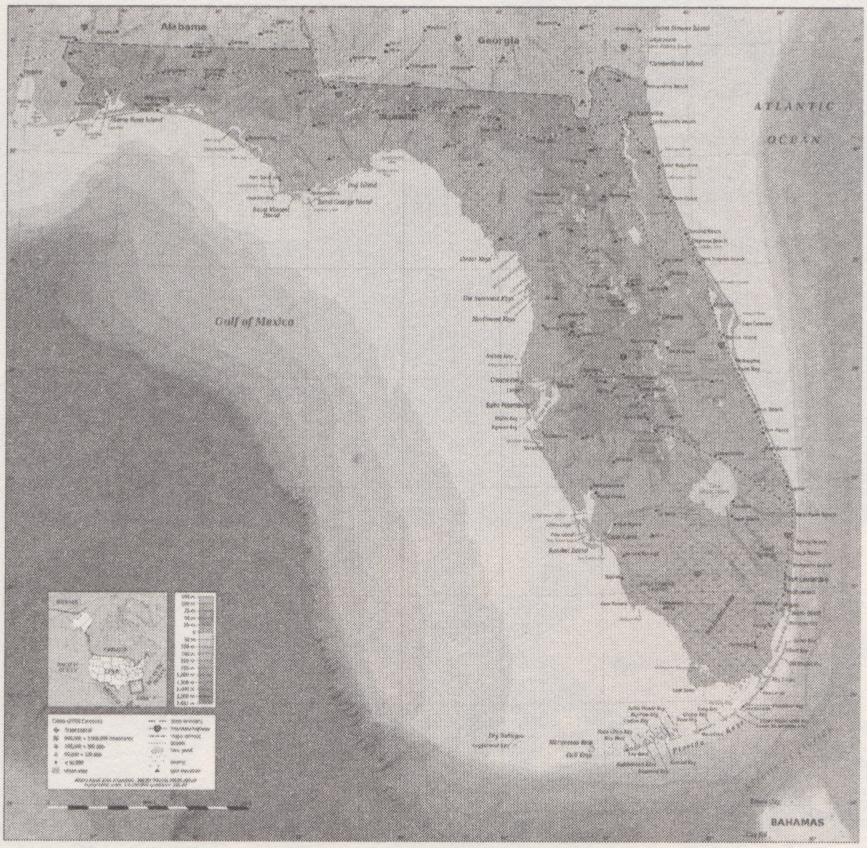
These properties tell us, furthermore, that the shape of a map created by conic projection is a circular arc, or an arc of a circular crown, and that the meridian and parallel grid is determined by the angular separation between the meridians and the way in which the parallels are spaced from one another. All of those details, together with the determination of which is or are the standard parallels, determine the map.



The shape of the grid of meridians and parallels is very characteristic in conic projections. This is Albers equal-area conic projection (1805), a good example.

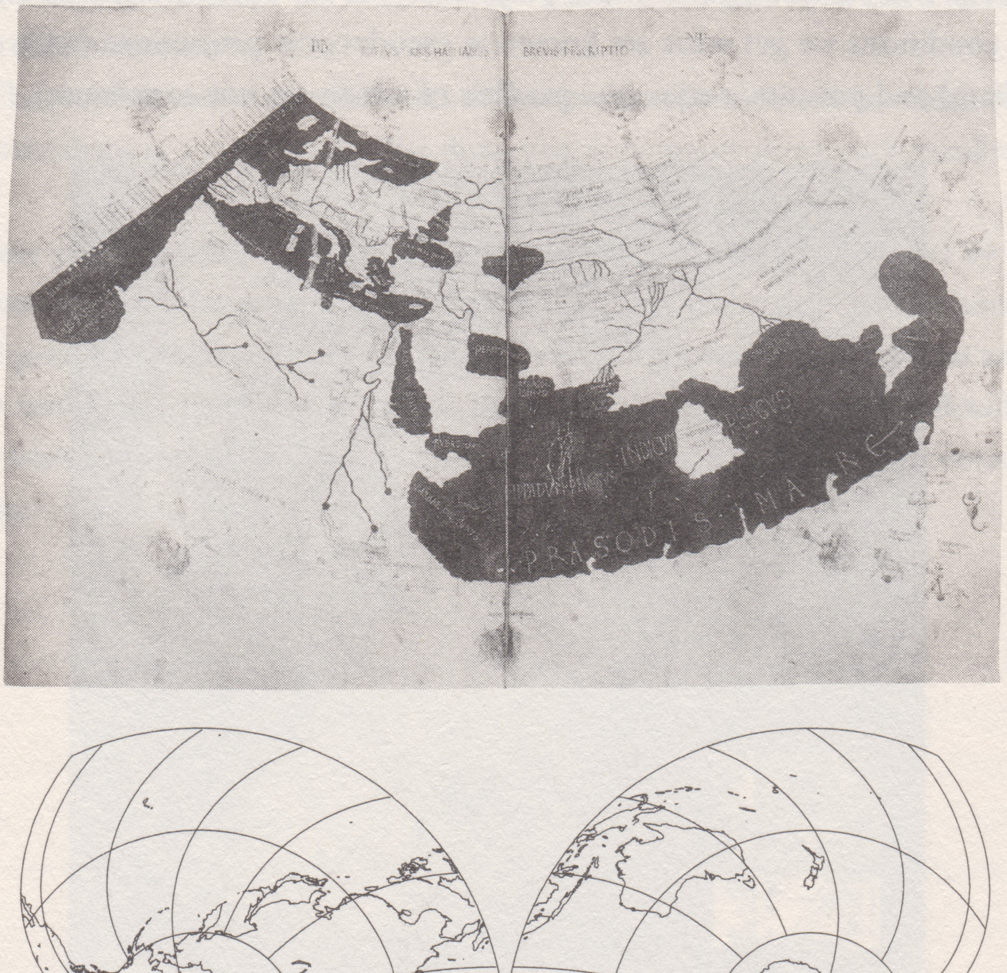
The distortion in maps designed from conic projections is small near the standard parallel or parallels, and increases in the direction of the poles. Due to this distortion, and the way in which these maps are made, they are usually used for regional or national areas or in somewhat more extensive territories in temperate zones. (In general, the azimuthal and cylindrical ones are used more in polar and equatorial regions, respectively.) They are particularly suitable projections for extensions of land situated between two meridians that are not very far from each other, such as Spain, France, Alaska or Mongolia, and can cover wider regions spreading in the east–west direction, such as Russia, Europe or the USA.

We are focussing on the normal, or polar aspect, but we could also consider the equatorial and oblique aspects of this projection. On the other hand, if we relax the conic conditions we get what are known as pseudoconic projections (with curved meridians) and polyconic ones (the parallels of which are not concentric).



A map of the Florida peninsula made by the Albers equal-area conic projection.

Ptolemy introduced two projections of the conic type (although he left no comments on the cone in the description he gave), both with concentric circumference arcs for the parallels, but one with the meridians represented by straight lines (illustration on page 126) and the other one by circular arcs (page 12). Ptolemy had an outstanding influence on cartography during the Renaissance and that's why from the beginning of the 16th century in particular we find the appearance of conic and pseudoconic projections which would be studied and used by the great cartographers such as Gerard and Rumold Mercator, Willem Blaeu, Joducus Hondius, Joseph de l'Isle and John Speed, among others. Some of them are very curious shapes, such as the heart-shaped projections by Johannes Werner and the French cartographer Rigobert Bonne (1727–1795) or the double-heart ones by the French mathematician and cartographer Oronce Finé (1494–1555).

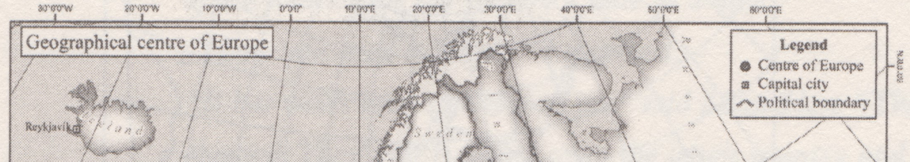


Above, a world map created by Ptolemy with a conic-type projection. Below, a map based on the double cordiform projection by Oronce Finé (1538).

The Lambert conformal conic projection

The cylinder and the plane can be seen as limit surfaces of the cone. We obtain the former when we move the vertex away and leave it in the infinite, and the latter if we bring it near to flatten the cone. Lambert used all the mathematical tools in his power (calculus, geometry, algebra and trigonometry) to develop a family of conformal conic projections, with two standard conformal parallels, with its limit cases being the stereographic, which is azimuthal, and the Mercator, which is cylindrical. As an example, Lambert included a map of Europe in his work.

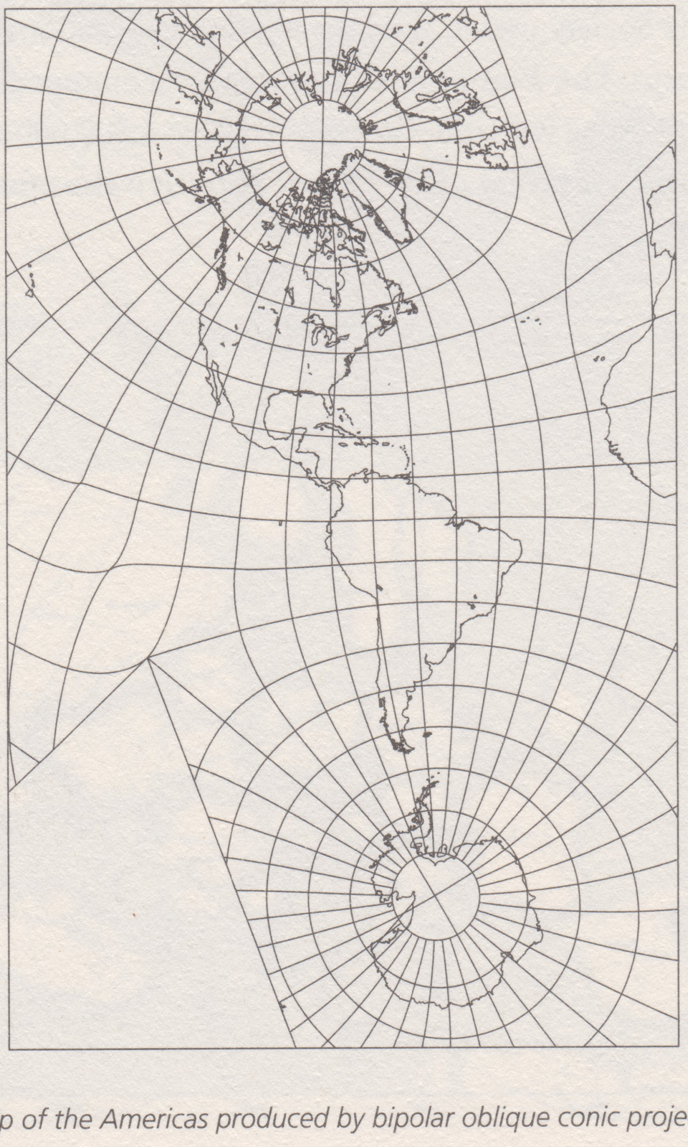
Until it was used by France during the First World War, this projection had remained largely forgotten. Since then it has become one of the most commonly used projections for 'large-scale' maps, after the Mercator projections. The United States Geographic Society uses it, as do many other international agencies for their topographical maps. The European Commission recommends it for conformal complete maps of Europe with scales lower or equal to 1:500,000 and it is commonly used in countries such as France or the United States. It is also often used for drawing up nautical charts.



Geographical centre of Europe
Legend
● Centre of Europe
▣ Capital city
Political boundary
Reykjavik
Atlantic Ocean
Dublin
London
Amsterdam
Brussels
Paris
Oslo
Stockholm
Copenhagen
Helsinki
Tallinn
Riga
Moscow
Berlin
Warsaw
Torun
Suchowola
Purnuškes
Kiev
Prague
Wien
Budapest
Dilove
Chisinau
Bern
Zagreb
Sarajevo
Belgrade
Bucaresti
Sofia
Madrid
Lisbon
Rome
Tirana
Skopje
Athens
Ankara
Nicosia
Beirut
Damascus
Amman
Jerusalem
Algiers
Tunis
Valletta
Rabat
Gibraltar
Mediterranean Sea
Tbilisi
Yerevan
Lambert conformal conic, ETRS 1989

A political map of Europe created by a Lambert conformal conic projection, which has become one of the most commonly used cartographic projections.

Other conic projections are the bipolar oblique, developed in 1941 by O.M. Miller and William A. Briesemeister of the American Geographic Society with the aim of creating a map to cover the whole of the American continent. The projection is still widely used for this purpose, and its design uses two oblique aspects of Lambert's conformal conic projection – the Albers equal-area conic projection, developed by the German cartographer Heinrich Ch. Albers in 1805, an equidistant conic projection which is reminiscent of Ptolemy's map; and the polyconic, usually attributed to the Swiss topographer Ferdinand R. Hassler (1770-1843), which uses different cones to operate and has a singular aspect looking like the curve named nephroide, in the shape of a kidney.



A map of the Americas produced by bipolar oblique conic projection.

Chapter 8

What Euler Told the Cartographer

> *"That's another thing we've learned from your nation," said Mein Herr, "mapmaking. But we've carried it much further than you. What do you consider the largest map that would be really useful?"*
>
> *"About six inches to the mile."*
>
> *"Only six inches!" exclaimed Mein Herr. "We very soon got to six yards to the mile. Then we tried a hundred yards to the mile. And then came the grandest idea of all! We actually made a map of the country, on the scale of a mile to a mile!"*
>
> *"Have you used it much?" I enquired.*
>
> *"It has never been spread out, yet," said Mein Herr. "The farmers objected. They said it would cover the whole country, and shut out the sunlight! So we now use the country itself, as its own map, and I assure you it does nearly as well."*
>
> Lewis Carroll, *Sylvie and Bruno* (vol. 2) [1893]

After our brief stroll around three significant examples of cartographic projections used for designing maps – Lambert's equal-area cylindrical projection plus the central and stereographic projections – which help give a better understanding of some aspects of the problems inherent to cartography, let's return to this book's central theme: do accurate maps of the Earth's surface really exist? How should such a map be built?

So as not to lose track of the purpose of our search, let's remind ourselves that a perfect map would be one that keeps all the metric properties, except for the change in size and scale, totally unchanged; that is, the area, the angles, the geodesics, the shapes and, in general, the lengths of the curves and distances. In other words, the cartographic projection to be used for its design must be isometric. Let's recall that a projection between surfaces is an isometry if the distances are preserved. That property implies, more specifically, that areas, angles and geodesics are preserved. With the aim of simplifying the search for and the construction of an accurate map we had asked ourselves whether being equal-area would be sufficient for the projection to be isometric, which would make our work a lot easier, because to design isometric maps, we could in that case limit ourselves to the construction of, or search for, projections that a priori simply preserve the area. We previously operated in an analogous fashion regarding angles and geodesics.

However, the three examples of cartography studied in the preceding chapters showed us that for a projection to be isometric and to obtain our ideal map, it is not enough to preserve just one of the three metric properties (area, angles and geodesics), but that it is necessary to add more information to the projection.

Equal-area conformal projections

After failing in our first attempt to obtain information that leads us to building the perfect map, what approach should we now take to the cartographic problem? Let's not make it too complicated for the moment and simply ask ourselves if complying with two of the three metric properties will be enough for it to be claimed that a projection is isometric.

We shall start by considering the case in which the projection of the sphere to the plane preserves the angles and the areas, and we'll see if it will necessarily be an isometry or not. To try and find an answer to this question we shall make use of what we studied in previous chapters on the distortion produced by projections that keep the area and the angles invariant. As we saw in Chapter 5, if the projection is conformal then the distortion does not vary with direction, and in particular, the distortion produced in the meridians (μ) is equal to the distortion produced in the parallels (λ):

$$\mu = \lambda.$$

On the other hand, in that same chapter it was also observed that, so as to compensate in area, in equal-area projections the deformation produced in meridians was inverse to that of the parallels, that is to say,

$$\mu = \frac{1}{\lambda}.$$

Taking into account both equalities we reach the conclusion that

$$\mu = \lambda = 1,$$

in other words, any projection that it is both equal-area and conformal will produce no distortion, either in the direction of the meridians and parallels or in any other direction; therefore it will be an isometry (the reader may wonder what happens to the scale, but remember that at the moment we are working with the reference sphere and, therefore, the change in size, i.e. the linear scale, does not form part of the problem).

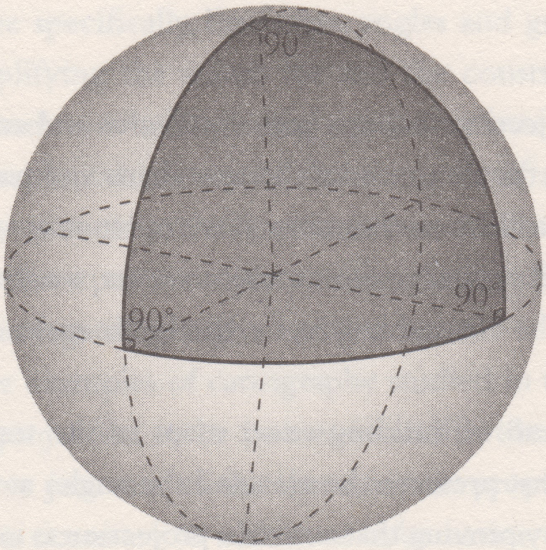
Eureka! One method of building exact maps of the terrestrial sphere could consist of projections that preserve the area and the angles at the same time, which is not a triviality, since by proving that a certain projection is isometric, all the metric properties are therefore preserved; in other words, all the geodesics, shapes, lengths of curves and distances.

Do accurate maps of the Earth exist?

Before attempting to find, or construct, an equal-area conformal transformation, from which a perfect map from the metric point of view could be constructed, let's continue with the study already begun and look at the case of those projections that preserve another two metric properties, such as angles and geodesics.

In a manner analogous to the concept of a triangle on the plane, that is to say, the region comprised by three straight lines which intersect two by two at three non-aligned points, the concept of 'spherical triangle' is defined as the region of the sphere bounded by three arcs of great circles that intersect two by two at three points that do not lie on the same great circle. As the projections that we are looking at preserve the geodesics, the image of a spherical triangle is a planar triangle, and as they are conformal, it preserves the angles of the triangles and also their sum. We know from classical geometry that 'the sum of the angles of a triangle is equal to π (180°)', but what do we know about the sum of the angles of a spherical triangle? Is it also equal to π (180°), as we might expect?

Let's begin with a specific example. Let's take a spherical triangle formed by a meridian arc between the North Pole and the Equator, another similar arc forming an angle of $\pi/2$ (90°) with the former, and the arc of the Equator that connects with the two former ones and that forms with each of them an angle of $\pi/2$ (90°), as shown in the diagram:



A spherical triangle formed by three angles of 90°, so their sum is 270°.

The sum of the angles of this spherical triangle is $3\pi/2$ (270°), and not π (180°), as we were expecting. In conclusion, there are no projections of the sphere to the plane that simultaneously preserve the angles and the geodesics. But this has a very important consequence, as it enables us to affirm there are no isometries of the sphere on the plane, in other words,

PERFECT MAPS DO NOT EXIST.

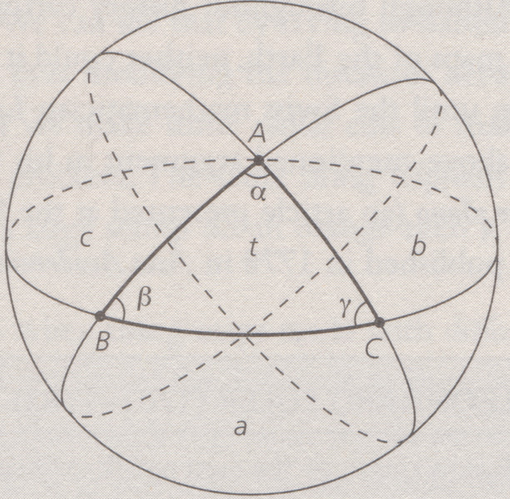
But, still further; this result is not only global, but also local. It is impossible to construct perfect maps of a small part of the Earth's surface. Put another way, it is impossible to construct local isometries of the sphere on a plane.

To prove it, let's go back to the sum of the angles of any spherical triangle. The value of this sum is between π and 3π (both values excluded). As each spherical angle is less than π, it is obvious that the sum of the three will be less than 3π but, furthermore, we can get as near to this value as we want. It is enough to imagine two of the vertices on the Equator and the other near to the Equator with the spherical triangle covering almost the whole of the hemisphere. At the other extreme we can consider two vertices on the Equator and the third one to the north in such a way that the meridian arcs are very close together, forming an angle as small as we want. The sum of the three angles of the triangle will be close to π. It could even be proven that for any spherical triangle this formula is fulfilled:

FORMULA OF THE SUM OF THE ANGLES OF A SPHERICAL TRIANGLE

Given a sphere of radius R, the part of it between two great circles (called a spherical lune) which intersect at an angle of α radians, has an area of $\alpha/2\pi$ times that of the whole the sphere, that is to say,

$$\frac{\alpha}{2\pi}\cdot(4\,\pi\,R^2).$$



Let a spherical triangle of vertices A, B, C have angles α, β and γ. If we take the great circles which contain $\overline{AB}$ and $\overline{AC}$, then the previous formula tells us that:

$$t+a=2\,\alpha R^2.$$

In a similar way we obtain:

$$t+b=2\beta R^2, \qquad t+c=2\,\gamma R^2.$$

By adding the previous equalities,

$$3t+a+b+c=2\,R^2\,(\alpha+\beta+\gamma).$$

But it turns out that $t+a+b+c$ is equal to the surface of a hemisphere (observe that for each vertex, for example A, there are two equal lunes of angle α, with two regions, each of areas t and a). In conclusion,

$$2t+2\,\pi\,R^2=2\,R^2\,(\alpha+\beta+\gamma),$$

and simplifying,

$$t=R^2\,(\alpha+\beta+\gamma-\pi).$$

$$\text{Spherical triangle area} = R^2\ (\text{Sum of the angles} - \pi),$$

if we consider R the radius of the sphere. As the sum of the angles of a spherical triangle of any size and shape is always greater than π, it is deduced that there are no projections of small regions of the sphere onto the plane that preserve angles and geodesics. Therefore, neither do local isometries exist. And the hope that we had of an equal-area conformal projection was an illusion, as no isometries of the sphere on the plane exist.

Although throughout history cartographers have not been able to make perfect maps of the Earth, neither could it be proven that it was not possible to make them until the Swiss mathematician Leonhard Euler proved it by making use of the above-mentioned argument in his work *De repraesentatione superficiei sphaericae super plano* (an article presented at the Saint Petersburg Science Academy in 1775 and published in 1778 in *Acta Academiae Scientiarum Imperialis Petropolitanae*).

LEONHARD EULER (1707–1783)

Euler is known as 'the most prolific mathematician of all time'. Throughout his life he published more than 500 books and articles, 866 if we add his posthumous works until 1911 when a start was made on editing his complete work, reckoned to amount to 90 volumes.

Euler was born in Basle. His father, a Calvinist pastor, wanted his son to study theology, but Leonhard preferred mathematics. At the age of 19, he published his first scientific paper on the optimum layout of ships' masts and sails without ever having seen a ship. Between 1727 and 1740 he spent his first period at the St. Petersburg Academy of Sciences. When he arrived there he found that the current tzar was not at all interested in science and, to get by, Euler had to sign up to the Russian navy for three years. He married Katherina Csell, with whom he became father to thirteen children! Euler used to say that he had made many of his discoveries with one or other of his children in his arms. It was round about that time that he lost the vision in his right eye. He spent the years between 1741 and 1766 at the Berlin Academy of Sciences.

Due to the economic crisis, he earned a living during his first years there by giving classes to the members of noble families. He did not have a good relationship with Frederick II, who used to call him "the Mathematical Cyclops", and despite being the head of mathematics, he had to take charge of tasks such as the levelling of the Finow Canal, constructing ornamental waterworks and water jets, the running of a salt mine and other diverse financial matters.

On his return to St. Petersburg, Catherine II treated him completely differently, both on a personal level and professionally. Euler spent his last years completely blind, but even so, nearly half of his work was written during that time.

Let's say it again: PERFECT MAPS DO NOT EXIST – of the Earth or of any part of it. Maps are always fallacious in some way, absolutely all of them. A couple of experiments demonstrate Euler's result. Let's take a plastic ball, cut it in half and try to place it completely flat on the table. We can see that by trying to flatten the ball it will become deformed or will tear, thus modifying the distances between the different points on the surface. Even if we make some radial cuts to make it easy to flatten we will still have the same problem. The same thing happens the other way round. If, for example, we stick a large stamp on an orange then it will form a lot of folds.

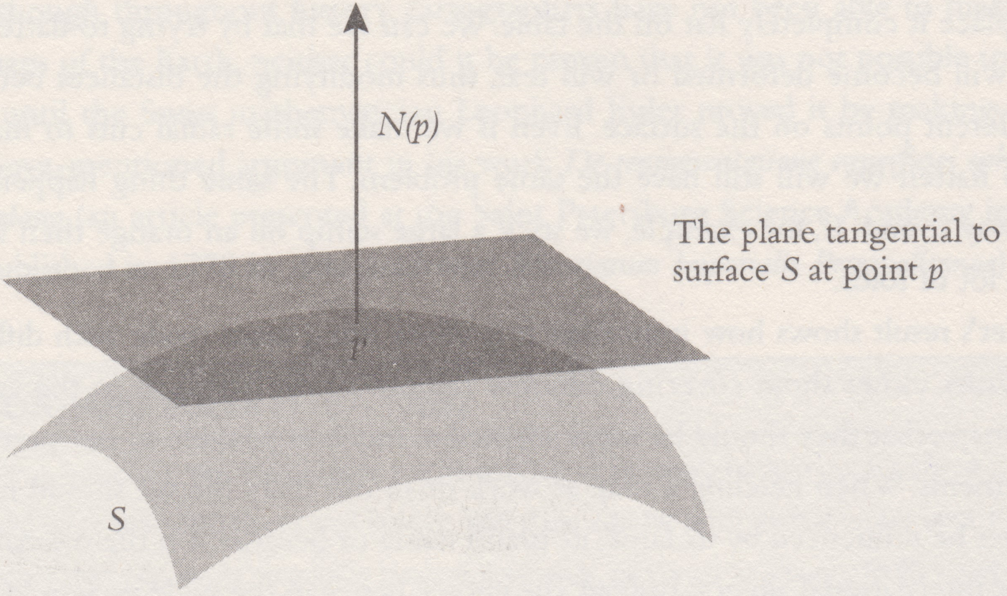
Euler's result shows how important it is when using maps made with different projections, either those covering different regions of the Earth or in the case of world maps, that they should be maps adapted as much as possible to our particular requirements. When needing a map at work or in our daily life we should not let ourselves be influenced by its fame, its brand name or because it is the official map of some international agency. Instead, the choice between the wide variety of maps available should be the consequence of an initial reflection on the properties we need to be preserved in it.

The Gauss curve, a return to the problem of cartography

One of the objectives that we had set ourselves at the beginning of this book, just the same as all cartographers have done throughout history, was the design of a correct terrestrial map. However, as we have proven by following in the steps of Euler, it is impossible. Leaving aside proofs, we can ask why such a map is not possible, why a transformation of the sphere to the plane that leaves its metric geometry invariant cannot be done.

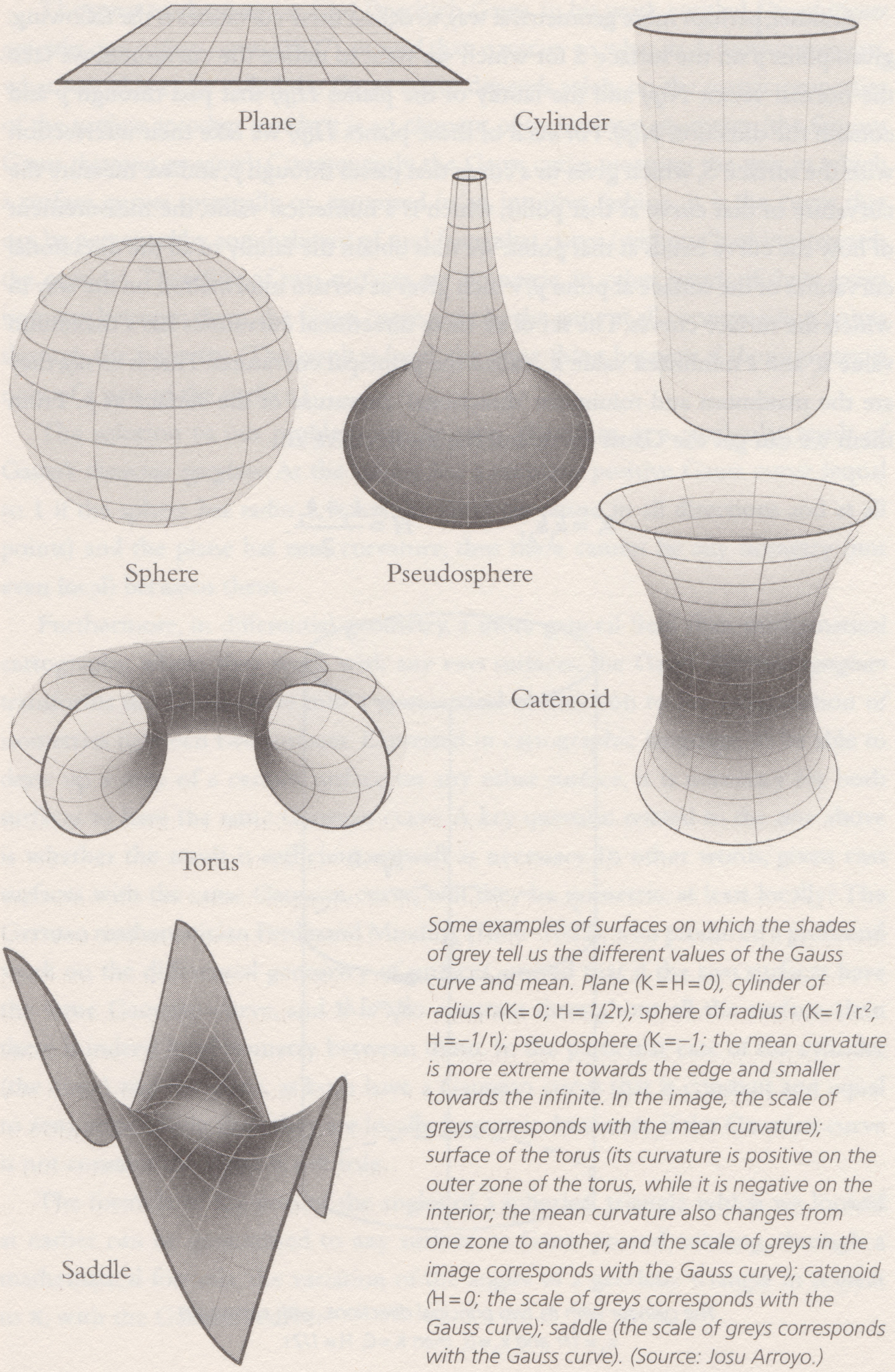
Quite possibly, bearing in mind the experiment with the plastic ball, the following answer will occur to us: "The sphere is a curved surface and the plane isn't!" And we would be right, bearing in mind what we have seen up to now, but it is also true that the cylinder and the cone are curved surfaces and, in spite of that, we can unroll them and leave them flat without altering their metric properties.

So what difference is there between the sphere and the other two surfaces, the cylinder and the cone? Is it possible that the way they curve is different, or could it be that the curvature problem is not as important as we thought? As we shall see next, not all surfaces are curved in the same way. How can we measure the way in which a surface is curved? The curvature at a certain point on the surface tells us how distant the surface is at that point, in its ambient space, from being a flat surface. But, somehow, we need to quantify it.



Two very important elements in the local study of a surface are the plane tangential to the surface at a point p *on the surface and the vector normal to the surface at* p, N(p), *which is perpendicular to the tangential plane.*

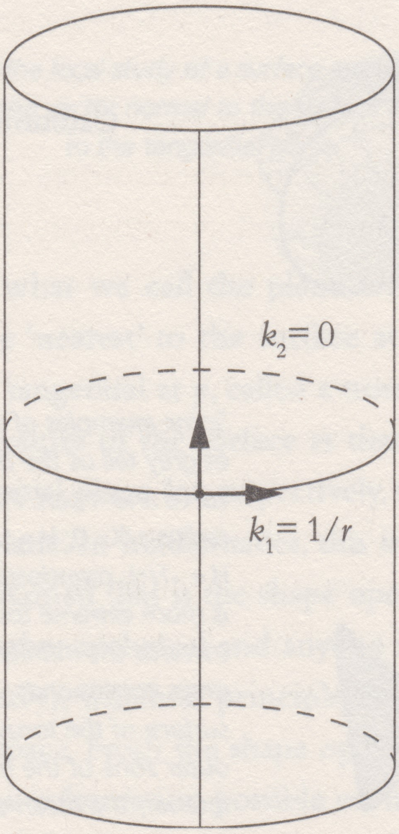
For this, let's consider what we call the plane tangential to the surface S at point p, which is the plane 'nearest' to the surface at that point, and the vector perpendicular to the plane tangential at p, called a normal vector (as shown in the diagram). To study the curvature of the surface at that point means to know, and to compute, how the tangential plane (or, respectively, the normal vector) varies in the ambient space of the plane. In mathematics, this is known as derivation. Thus a mathematical object is obtained that is the shape operator (we shan't go into the specific definition as it is rather technical and anyone interested can find it in any book on differential geometry), which in principle contains all the information of how the surface curves in space. From the shape operator two different curvatures are defined to extract all the information possible on it, the curve called the Gauss curve K and the mean curvature H.



Some examples of surfaces on which the shades of grey tell us the different values of the Gauss curve and mean. Plane ($K = H = 0$), cylinder of radius r ($K = 0$; $H = 1/2r$); sphere of radius r ($K = 1/r^2$, $H = -1/r$); pseudosphere ($K = -1$; the mean curvature is more extreme towards the edge and smaller towards the infinite. In the image, the scale of greys corresponds with the mean curvature); surface of the torus (its curvature is positive on the outer zone of the torus, while it is negative on the interior; the mean curvature also changes from one zone to another, and the scale of greys in the image corresponds with the Gauss curve); catenoid ($H = 0$; the scale of greys corresponds with the Gauss curve); saddle (the scale of greys corresponds with the Gauss curve). (Source: Josu Arroyo.)

Another, perhaps more geometrical way, to define these curvatures is the following: given point p on the surface S for which we want to define the curvatures, we take the normal vector $N(p)$ and the family of the planes $\Pi(p)$ that pass through p and contain the direction $N(p)$. For each of those planes $\Pi(p)$ we take their intersection with the surface S, which gives us a curve that passes through p, and we measure the curvature of that curve at that point, which is a numerical value, the measurement of how the curve bends at that point. We thus obtain the family of all the directional curvatures of the surface at point p, which gives us certain information on the way in which the surface curves. The set of all these directional curvatures has a maximum value k_1 and a minimum value k_2, called the principal curvatures. That is to say, they are the maximum and minimum 'directional' curvature of the surface at p. From them we can get the Gauss curve and the mean curve as:

$$K = k_1 k_2, \qquad H = \frac{k_1 + k_2}{2}.$$



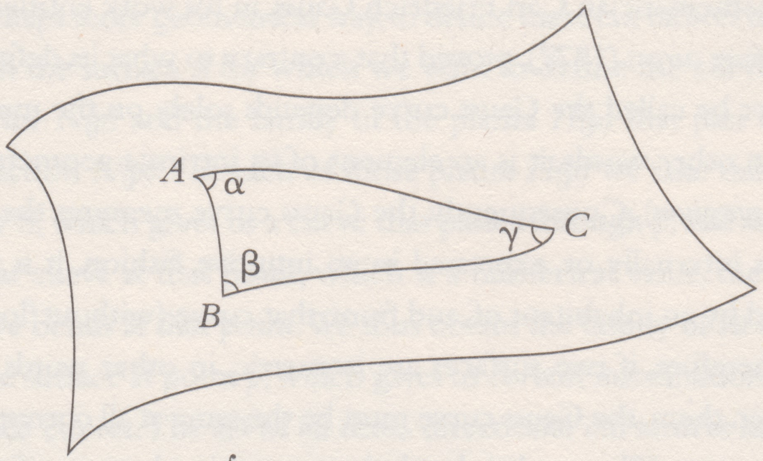
The cylinder with its two principal directions, with curvatures $k_1 = 1/r$ *and* $k_2 = 0$, *then* $K = 0$, $H = 1/2r$.

The great mathematician Carl Friedrich Gauss, in his work entitled *Disquisitiones generales circa surfaces curvas* (1827), proved that, contrary to what its definition suggests, what would later be called the Gauss curve depends solely on the metric properties of the surface, in other words, it is an element of its intrinsic geometry (the famous Gauss *theorema egregium*). Consequently, the Gauss curve measures the way in which a surface curves internally or, expressed in an intuitive fashion, it is the curve that can be perceived by an inhabitant of, and from, that curve (without 'looking' towards the outside). Therefore, if two surfaces are isometric, in other words, if there exists isometry between them, the Gauss curve must be the same at all corresponding points through the isometry. (The result is local; the same thing happens if the isometry is only of a part of the surface.)

The solution to our problem can be seen, therefore, as a particular result of Gauss's *theorema egregium*. As the sphere has a constant positive Gauss curve (equal to 1 if the sphere has radio 1, it has the same curvature in all directions and at all points) and the plane has zero curvature; thus there cannot be any isometry (not even local) between them.

Furthermore, in differential geometry, a more general field than mathematical cartography, where they work with any two surfaces, the Gauss *theorema egregium* transforms the Gauss curve into a geometrical obstruction to the construction of isometries between two surfaces. Expressed in cartographic language, to be able to draw up a map of a certain surface on any other surface, it is necessary for both surfaces to have the same Gaussian curve. A key question related to the one above is whether the result is sufficient, as well as necessary. In other words, given two surfaces with the same Gaussian curve, will they be isometric, at least locally? The German mathematician Ferdinand Minding (1806–1885), who produced a profound work on the differential geometry of surfaces, proved that if the two surfaces have the same Gaussian curve, and it is also constant throughout all the surface, then there is indeed local isometry between them. In the particular case of the cylinder (the cone) and the plane, as both have a Gaussian curve that is constant and equal to zero, respectively, then they are locally isometric. However, if the Gaussian curve is not constant the result is not true.

The formula of the sum of the angles of a spherical triangle which we looked at earlier can be generalised to any surface, as Gauss proved, relating, through a mathematical formula, the variation of the angles of a geodesic triangle in respect to π, with the Gaussian curve:



$$\int_T K\,d\sigma = \alpha+\beta+\gamma-\pi.$$

The formula of the sum of the angles of a triangle, both on the plane and on a sphere, can be generalised to any surface; this is the well-known Gauss–Bonnet formula, which refers to Gaussian surface curvature.

Naturally, in the case of the sphere of radius R we refer back to the previous formula, as the Gaussian curvature is $1/R^2$, and thus:

$$\int_T K d\sigma = \frac{1}{R^2}\int_T 1\ d\sigma = \frac{1}{R^2}\,\textit{Area of } T.$$

But we have introduced two curvatures; therefore, we can ask ourselves what happens with the other one, the mean curvature. The extrinsic geometry of a surface determines the relationship of the surface with the exterior three-dimensional space, and the extent to which it really curves in that space is determined by the mean curvature.

The terrestrial globe

It is not possible to design accurate maps of the Earth. The terrestrial globe is the most perfect representation of our planet and preserves all the metric properties we looked at above, except for the size. The only distortion produced is the scale factor, which is constant all over the globe. On this model of the Earth we can trace accurate flight and maritime routes as the directions and distances are authentic. The distance between two points on the terrestrial surface, such as two cities, for example, can be calculated by tracing the great circle on the globe, which can be found easily

THE HISTORY OF THE TERRESTRIAL GLOBE

The Greeks, who were aware of the Earth's spherical shape, had already begun to construct terrestrial globes, but it was to be the grammarian and Stoic philosopher Crates of Mallus who was to construct the first recorded globe. That was around 150 BC, before the discoveries of America, Australia and part of the African continent, and the globe showed four terrestrial zones spaced out around the world, though only one was known, the ecumene. The Greeks, Romans, and later the Arabs, created and used both terrestrial and celestial globes.

The oldest terrestrial globe that has been preserved was constructed by the German geographer Martin Behaim in 1492. The Renaissance was a marvellous epoch in the construction of these cartographic tools, and German cartographer Martin Waldseemüller (c. 1470–c. 1520) made an important breakthrough in the mass production of globes by using printed spherical gores (triangle shapes).

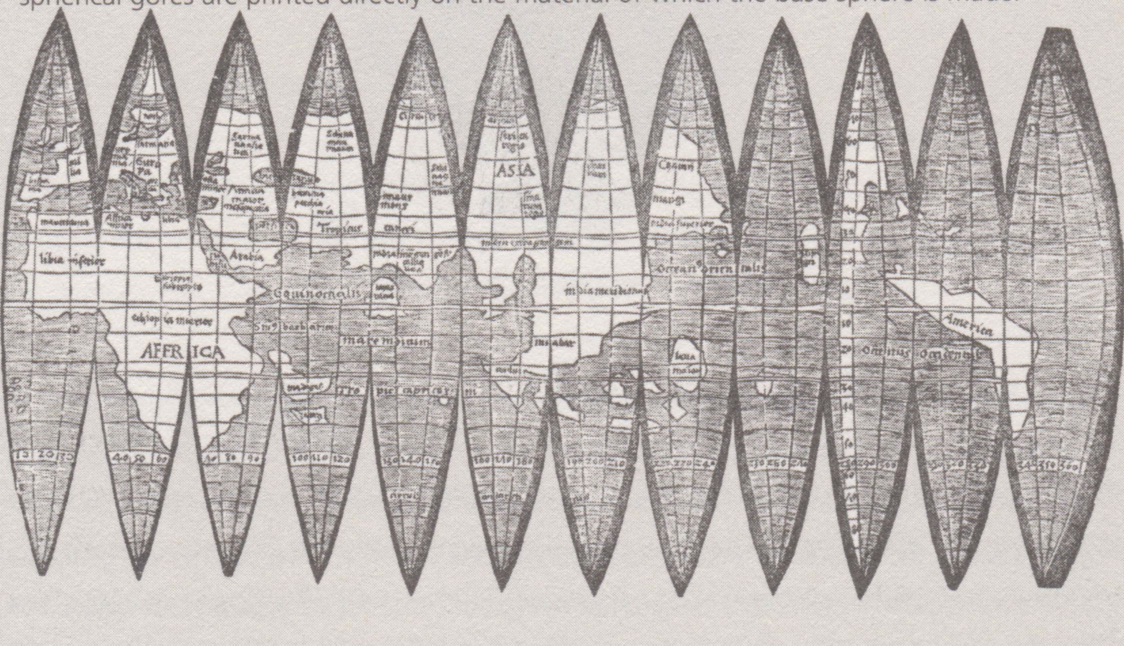
A facsimile of Waldseemüller's globe (1507).

By studying terrestrial globes throughout history we can see the parallel developments in their construction and in geographic knowledge. The Flemish cartographer Gerardus Mercator was responsible for a turning point both in the technical development of the making of these globes and in the scientific aspects associated with the cartographic problems. His goal was to make the globe a useful tool in the art of navigation, both for scholars and for sailors. For example, he included images of the loxodromic curves. However, many of his pieces simply ended up as handsome ornaments in the homes of the upper classes.

HOW IS A TERRESTRIAL GLOBE BUILT?

Despite the fact that, basically, only a simple sphere is needed to recreate the Earth correctly, its actual construction entails very many technical problems. The first one is the size. Terrestrial globes are normally 30 to 40 cm in diameter, which makes them too small for details to be seen. For example, if a globe had relief to scale, Mount Everest would be about 0.28 mm high, and the course of a river would be imperceptible. The second problem is the material used to make the base sphere. Traditionally, globes were solid and might be made out of glass, marble, wood or metal (gold, silver, bronze or lead), but after Mercator they became hollow, and consisted, for example, of a wooden frame onto which a paste of paper and plaster was applied. Nowadays they continue to be made hollow, but out of diverse materials such as cardboard, plastic or metal.

Since the days of Waldseemüller, a series of gores or spherical lunes have been printed and then stuck onto the base sphere to create globes. We therefore come up against the same problem as that for creating maps, as the image of the terrestrial sphere has to be printed on the planar gores. Normally, 12 spherical gores centred on the Equator are used, with the two polar regions done separately, designed with a modified sinusoidal projection. A more frequent alternative nowadays is to place two sets of 12 triangular sectors centred on one of the poles and covering one of the hemispheres. However, with modern techniques, the spherical gores are printed directly on the material of which the base sphere is made.



Spherical gores from Martin Waldseemüller's globe (1507).

using a string pulled tight and then by measuring the circle with a tape measure, or by measuring the string itself, and then by applying the scale to the measurement obtained. Likewise, other measurements of interest can be taken by using a terrestrial globe, and will be more exact than those of any planar map. The errors that may be obtained in measurements on the terrestrial globe will be a result of errors made in the process of measuring – not from the globe itself, provided it has been correctly constructed. (As we shall see, the complexity involved in the construction of globes means that it is in the actual building of them where errors originate.)

A modern terrestrial globe.

Terrestrial globes are very useful in cartography, geography, marine navigation, geodesics, oceanography, climatology and seismology, among other sciences. But furthermore, they provide a true image of the Earth, of its roundness and its corresponding spatial relationships. For that reason it would be useful for all schools and centres of education and every home to have a terrestrial globe to enable us to observe a true image of our planet. At the same time, the frame on which they are fitted keeps us aware of the terrestrial rotation, the part we can see would be the part of the Earth where it is daytime, that is, the part illuminated by the Sun, while the part we cannot see would be the region where it is nighttime.

Despite the theoretical perfection of terrestrial globes, their unassailable disadvantages rule out their use (even assuming that they have been constructed perfectly):

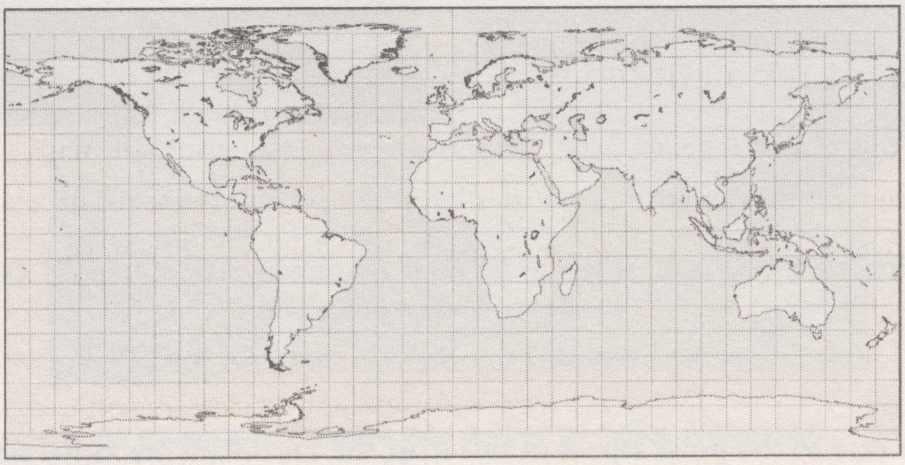
1. They are fragile, bulky and difficult to handle, transport and store.
2. They are expensive to produce, particularly the larger ones, and not at all practical for showing details.
3. They are difficult to handle when taking measurements or fixing angles.
4. Only one hemisphere can be worked with at a time.
5. They are totally impracticable for making reproductions either by printing or by electronic means.

Equidistant projections

To finish this chapter we are going to introduce a new family of maps classified in accordance with their metric properties. Throughout this book we have emphasised that the ideal for any cartographer would be to design a map with constant scale (reduction index), in other words, with no other distortion other than the change in size. We have just proved that this is not possible, as the scale of any planar representation of the Earth is not constant since it varies both in position and in direction due to the fact that cartographic projections inevitably generate distortion. However, there is a way to construct maps on which a family of curves has constant scale and lengths proportional to the corresponding curves on the Earth (these curves are called standard lines or curves). Such projections are called equidistant. Let's look at three examples of this type of projection, one from each type: cylindrical, azimuthal and conic.

Equidistant cylindrical projection

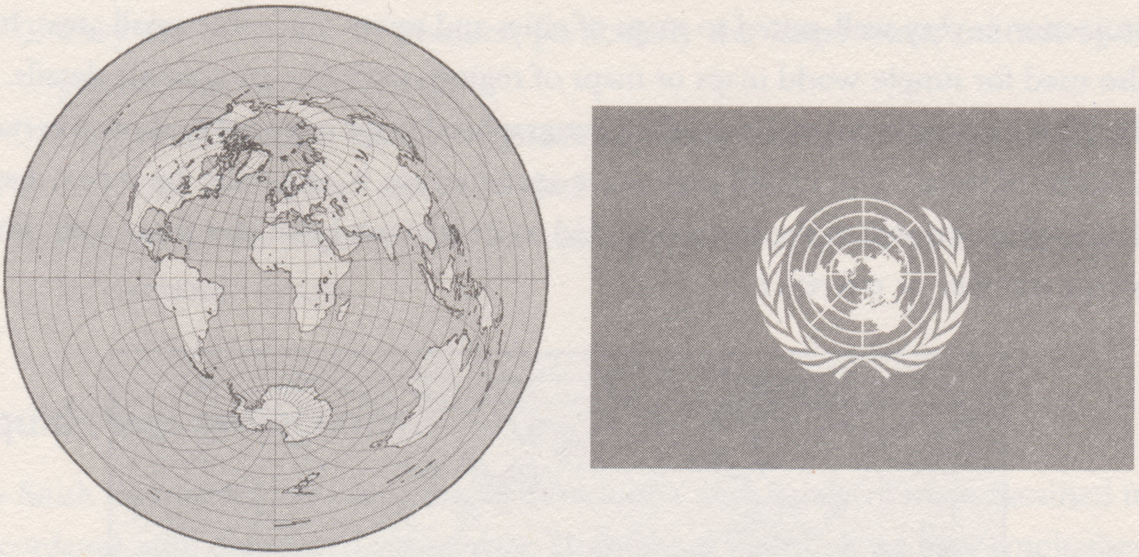
This projection, also known as rectangular, equi-rectangular or planar chart, is mathematically trivial and, in its simplest case and with the Equator as a tangent curve, defined by taking the longitude and latitude directly as Cartesian coordinates (next figure). While Lambert's equal-area cylindrical map becomes compressed at high latitudes and Mercator's expands, on the equidistant cylindrical map the parallels are equally spaced. The distance from one parallel to the next, on the map, is always the same. The scale is correct along the meridians and the Equator (in this case the grid is formed by squares – *plate carrée* or planar chart), or also along the meridians and a pair of parallels equidistant from the Equator (equi-rectangular or rectangular). This projection is usually attributed to Eratosthenes, though Ptolemy gives Marinus of Tyre as its inventor around the year 100. From then onwards it was widely used, particularly in marine navigation, due to the simplicity of its construction. This projection is very well-suited to maps of cities and towns and other small areas. It is also used for simple world maps or maps of regions with few geographic details. In the 20th century it was rarely used for geographical maps required to show a certain amount of detail. The USGS and other agencies use it, as a rule, for index maps, that is, those on which the maps included in an atlas are listed and those indicating a page number or reference.



A map produced with the 'planar chart' projection, which is a particular case of the equi-rectangular projection with standard parallel on the Equator (source: Carlos A. Furuti).

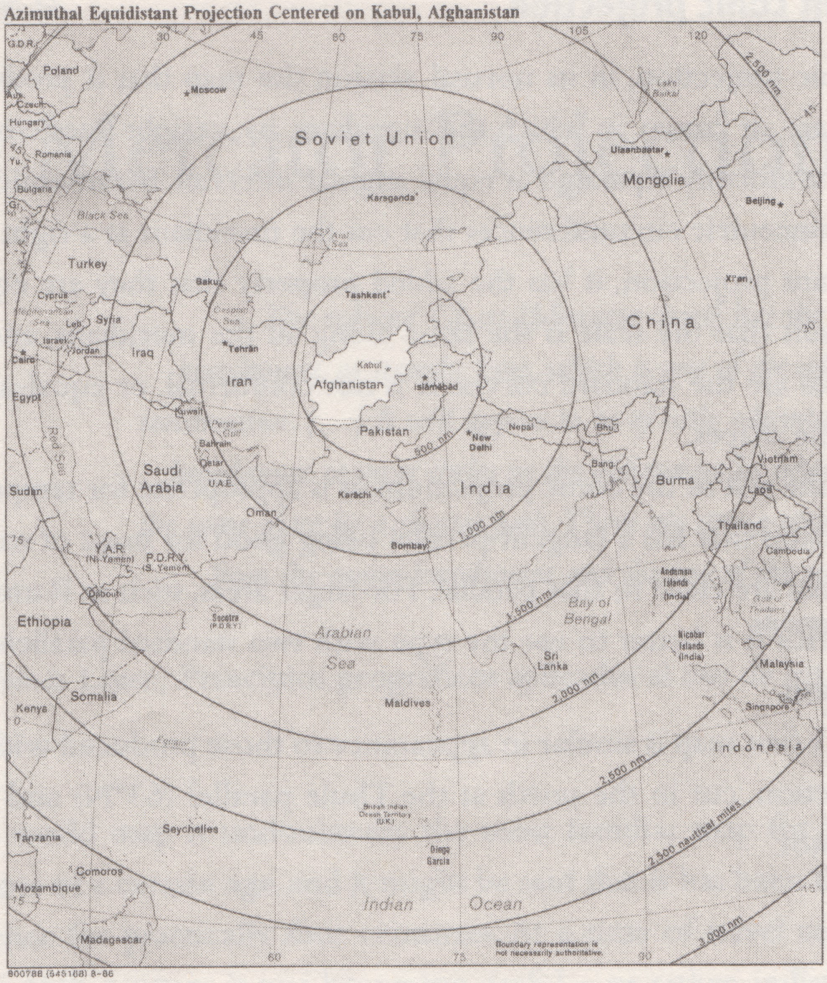
Equidistant azimuthal projection

This is the fourth classical azimuthal projection, but unlike the other three mentioned earlier, it is not geometric. Like other azimuthal projections, it fulfils the property of having the geodesics – the great circles – that pass through the reference point transformed on the plane into straight lines passing through the central point, and also having the angle between those geodesics preserved. This projection's special property is that the scale is constant along the straight lines passing through the centre of the map (they are the standard lines of that equidistant projection). In other words, the distances from the centre of the map are preserved. Furthermore, the Earth can be presented as a whole on one map, although the distortion becomes very big on crossing the great circle bounding the hemisphere that has the extreme reference point, and the limit of the map is the antipodal point that corresponds with all the exterior circumferences. On this map there is one particular point, which becomes the 'centre of the world', and from which distances are preserved, for which reason is normally called an 'egocentric map'.



An equidistant azimuthal map centred on the North Pole (source: Carlos A. Furuti). On the right, the flag or emblem of the UN.

In the case of the polar aspect of this map, the meridians are radial straight lines emanating from the centre, i.e. the image of the reference point, and the parallels are equally spaced concentric circumferences. The map centred on the North Pole is very familiar to us all as it forms part of the flag and emblem of the United Nations Organisation, the UN. It shows the Antarctic replaced by olive branches. In its polar aspect, the simplicity of this projection leads us to suspect that it has been used since very early times. The Egyptians are thought to have used it to draw stellar maps, though the oldest known celestial map for which it was used was produced by Conrad de Dyffenbach in 1426. However, the first to use it for making terrestrial maps was Mercator for his famous world map in 1569, which included two discs for the polar regions produced with this projection. From then onwards it was widely used. For example, it appeared in a large number of atlases of the polar regions in the last two centuries, or together with world maps showing those regions, such as in the map made with the Van der Gritten projection for the National Geographic Society or the USGS's maps. It also came to be widely used for maps of one hemisphere, north or south, and even for all the planet.


Azimuthal Equidistant Projection Centered on Kabul, Afghanistan
Soviet Union
Moscow
Poland
Czech.
Hungary
Romania
Yu.
Bulgaria
Black Sea
Turkey
Cyprus
Syria
Leb.
Israel
Jordan
Cairo
Egypt
Iraq
Baku
Caspian Sea
Tehrān
Iran
Kuwait
Persian Gulf
Bahrain
Qatar
U.A.E.
Saudi Arabia
Red Sea
Sudan
Y.A.R. (N. Yemen)
P.D.R.Y. (S. Yemen)
Djibouti
Ethiopia
Somalia
Kenya
Equator
Tanzania
Seychelles
Comoros
Mozambique
Madagascar
Karaganda
Aral Sea
Tashkent
Kabul
Afghanistan
Islāmābād
Pakistan
New Delhi
Nepal
Bhutan
Bang.
Karāchi
Oman
India
Bombay
Socotra (P.D.R.Y.)
Arabian Sea
Maldives
British Indian Ocean Territory (U.K.)
Diego Garcia
Indian Ocean
Sri Lanka
Bay of Bengal
Andaman Islands (India)
Nicobar Islands (India)
Burma
Thailand
Laos
Vietnam
Cambodia
Gulf of Thailand
Malaysia
Singapore
Indonesia
China
Xi'an
Beijing
Mongolia
Ulaanbaatar
Lake Baikal
500 nm
1,000 nm
1,500 nm
2,000 nm
2,500 nm
2,500 nautical miles
3,000 nm
30
45
60
75
90
105
120
15
Boundary representation is not necessarily authoritative.
800788 (545168) 8-86

An equidistant azimuthal map centred on Kabul, Afghanistan.

The difficulty in constructing the oblique and equatorial aspects of this projection meant that it was not used until the 19th century for those aspects. The oblique aspect has been used for world maps centred on large cities and important places, and sometimes also for maps of continents. This is the opposite of what happens with the equatorial aspect, perhaps because the Equator does not cover many places considered 'important' by mapmakers.

This is a projection that is very useful for cases in which it is necessary to focus on distances or the shortest paths from a particular point. For instance, it might be used by a member of the military in charge of a nuclear base for determining which cities are within their strike radius, or in marine and air navigation to establish a fixed bearing from a port or airport to different places in the world, or, in combination with the Mercator projection (as will be covered in Chapter 9), to navigate between two spots. Besides being useful for navigation, it is also useful in research on earthquakes. Radio operators using directional antennae also use it for marking the direction and amplifying the signal.

Equidistant conic projection

Being a conic projection, in its normal version the map that it produces is in the shape of an arc, or crown, in which the meridians are straight lines emanating from a point and uniformly separated at equal angular intervals. Meanwhile, the parallels are arcs of concentric circumferences that cut the meridians at a right angle but, in the equidistant projection, it has the added property that they are equally spaced, with the result that the scale is the same along all the meridians, and means that the meridians are the standard curves. It is not conformal, or equal-area, nor does it preserve the shapes.

As with the rest of the conic projections, it is appropriate for temperate regions, with the version that has a tangent parallel being good for maps of small countries or regions that are close to that parallel. For larger areas, such as Russia, Europe or North America it is better to use the type with two standard parallels crossing the relevant region.

Ptolemy's first map is similar in appearance to those produced with equidistant conic projection, cut in the north at the Thule parallel (63°N) and in the south by the Equator.

Chapter 9

The Mercator Projection

> *The reduced Chart [Mercator chart] has the rhumb lines, or loxodromes, as straight lines, which is one of the advantages of this construction. [...] It will appear to be of little advantage to follow the loxodromes or great circles in the oblique paths when you can reach the same place by a shorter path. There are strong reasons not to stray from the major loxodromes and the use of the compass, which is always the best way...*
>
> *Tomás López,* Principios geográficos aplicados al uso de los mapas *(1783)*

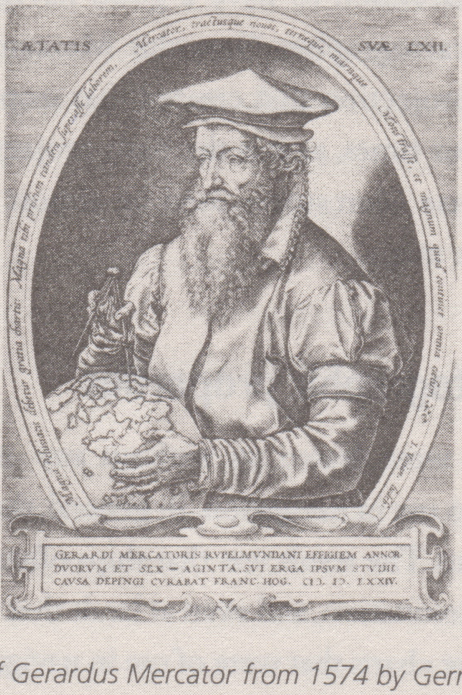
Mercator's world map is undoubtedly the most familiar map for most of us, at least for those of a certain age, and it might be said that it has been the world map par excellence for practically four centuries. At a time of travel and discoveries, back in the 16th century, it was becoming useful – or rather, indispensable – for mariners, traders and statesmen to have a good navigation map available. The Flemish scientists and cartographer Gerardus Mercator would develop such a map. Furthermore, the projection that created that map has surely been the most valuable and useful of our times, as it is the basis projection for the UTM system – the Universal Transverse Mercator – used by almost all international agencies for designing large-scale maps.

Definition and cartographic properties

Mediaeval maps, which were drawn up without any mathematical projection or scientific basis, were of no use in navigation. In general, they were no good for taking any kind of measurements. The result was that ships often arrived at places very distant from their marked destination, and even at lands that were previously unknown. The first attempts to make charts that were of use in navigation came from those who needed them most, the mariners, whose portolan charts were based on notes they took on their voyages, on their observations of the stars, and on their measurements

GERARDUS MERCATOR (1512–1594)

Mercator was a brilliant man who dedicated himself both to applied knowledge – it could be said in this regard that he was a cartographer, a geographer, a calligrapher, an engraver, a builder of scientific instruments and a publisher – and to theory. He had a great interest in mathematics, astronomy, cosmography, terrestrial magnetism, history, philosophy and theology. His outstanding map-making achievements include the 'map of Palestine' of 1537, which clearly shows a religious motivation; his first world map, produced with an Oronce Finé double-heart projection – and the first to use the word 'America' to refer to the New World; a famous 1554 map of Europe, created by using a Werner projection; and the 'Mercator map' of 1569; in addition to his 1541 terrestrial globe, which was then the most complete globe ever made and even included loxodromes. His final project was to draw up a map of the world by producing individual maps of the different regions. He published the first part of this atlas (he was also the first to use the term 'atlas' to refer to a collection of maps, in honour of the Titan, Atlas, the god of astronomy who was condemned by Zeus to hold up the heavens on his shoulders) in 1585, with 51 maps, centred on Germany, France and the Netherlands. For the 1589 volume he would add 23 maps of Italy and Greece. His son Rumold published the Mercator Atlas in 1595 with 28 more maps showing different parts of Europe. Various cartographic projections were used – conic, stereographic, the Sanson-Flamsteed, Werner's, the Mercator and several others.


ÆTATIS
SVÆ LXII.
GERARDI MERCATORIS RVPELMVNDANI EFFIGIEM ANNOR
DVORVM ET SEX – AGINTA. SVI ERGA IPSVM STVDII
CAVSA DEPINGI CVRABAT FRANC. HOG. CIƆ. IƆ. LXXIV.

A portrait of Gerardus Mercator from 1574 by German painter Frans Hogenberg (1535–1590).

of distances and directions. The instruments they used were a compass, a half-circle protractor marked in degrees, a ruler and a pair of compasses. Portolan charts, however, did not take the sphere's geometry into account, that is, its curvature or roundness, and no projection whatsoever was used in designing them.

The first scientific attempts to solve the problem of the nautical charts came from the hands of Mercator, Ortelius and other cartographers of the time, who put science to use in drawing up maps and charts. Gerardus Mercator's objective was, therefore, to construct a world map that was useful for navigation, which meant it had to preserve directions (in other words, the projection would have to be conformal), while the loxodromic curves, the rhumb lines, would have to be shown as straight lines.

The Portuguese astronomer and mathematician Pedro Nunes (1502-1578) had already studied and described the loxodromic curves (which are spirals on the terrestrial surface that converge on the poles) in his *Treatise Defending the Sea Chart* (1537). In it, he rejected the widely held belief that sailing in a fixed direction described a great circle, that is the curve of minimal distance. To sail between two points on the Earth they tried to follow the shortest path, the orthodromic. However, that meant that they had to constantly change direction, which made it easy to deviate from the route. Consequently, it was better to sail on the loxodromic route, as with this system it was enough to maintain a constant bearing, even though the path was longer. Mercator had already included a multitude of loxodromic curves on the terrestrial globe of 1541.

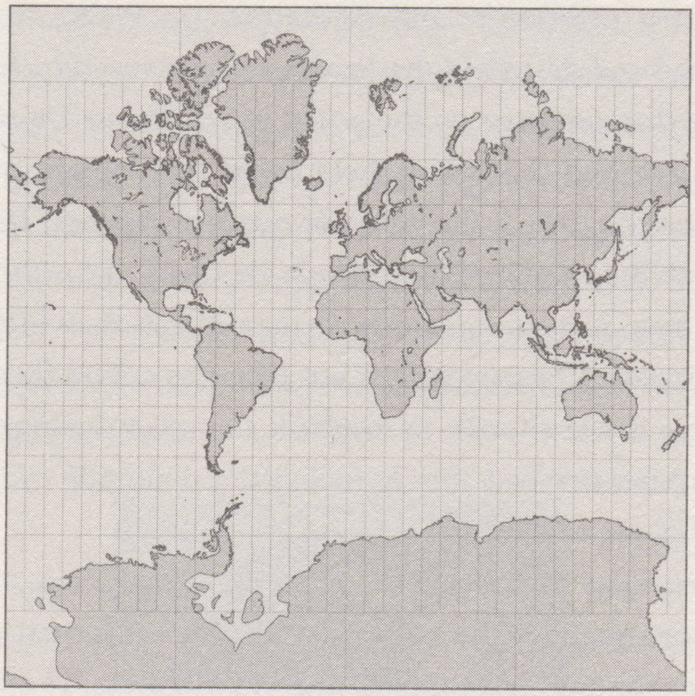
The geometric issue underlying the problem of designing a nautical chart like the one Mercator sought is the construction of a conformal projection, which transforms the loxodromes into straight lines on the plane. To start with, on a map such as this one, the meridians and parallels have to be represented by straight lines that intersect perpendicularly.

What's more, from the study carried out in Chapter 5 on the Lambert projection we know that this cylindrical projection is not conformal because the distortion produced in the direction of the meridians, $\cos\phi$, is not equal to that produced in the direction of the parallels, $1/\cos\phi = \sec\phi$ (ϕ is the latitude of the point where we are). So, by intuition, we see that the solution lies in managing to modify the map so that both distortions coincide. More specifically, we have to 'stretch' Lambert's map in the north-south direction, as the map goes from being shrunken (the distortion on the parallels is $\cos\phi$) to being stretched (new distortion equal to $1/\cos\phi = \sec\phi$). That is the key idea behind the construction, and if we now integrate this distortion mathematically, we shall get the required projection's mathematical expression, the Mercator projection:

$$x = R\left(\theta - \theta_0\right)$$

$$y = R \ln\left[\tan\left(\frac{\pi}{4} + \frac{\varphi}{2}\right)\right],$$

with θ being longitude (θ_0, the central longitude of the map) and ϕ, the latitude, and with the reference sphere, $R = 1$.



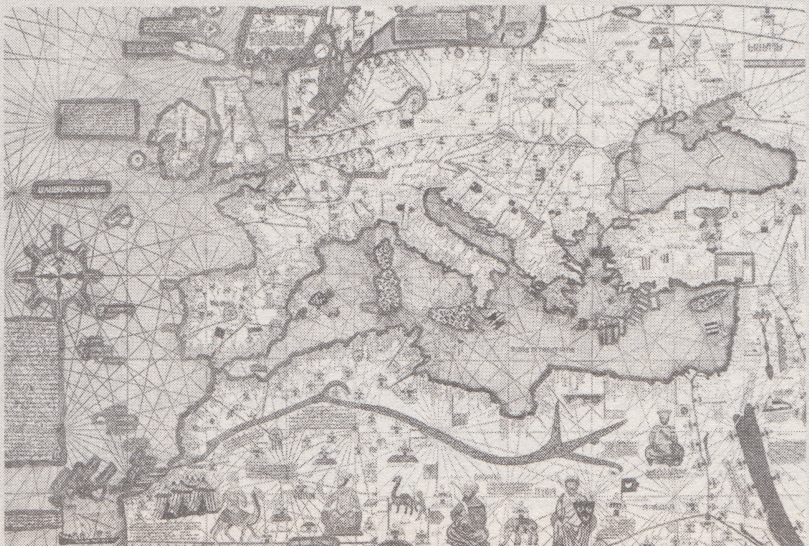
A modern map created by the Mercator projection (source: Carlos A. Furuti).

This is essentially what Mercator did on his map *Nova et aucta orbis terrae descriptio ad usum navigatum emendate accommodata* (*A New and Extended Description of the Earth with Corrections for Use in Navigation*) in 1569. He built a grid of meridians and parallels represented by straight lines that intersect perpendicularly, and then to space the parallels in such a way as to compensate for the distortion produced in the meridians and thus achieve the same distortion in both directions, as can be seen on the original map from 1569.

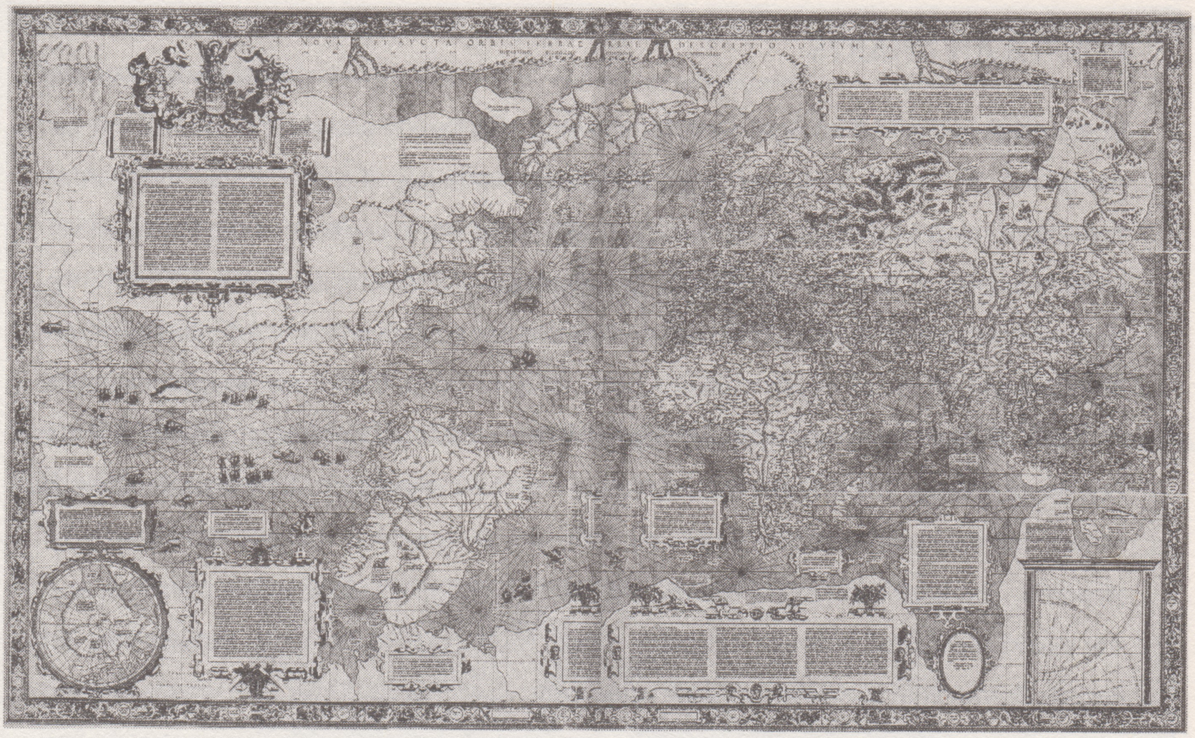
PORTOLAN CHARTS

World maps in the Middle Ages were for showing information and were of no real use for navigation. To voyage from port to port, mariners made use of their own personal notes based on distances calculated, astronomical observations and knowledge of the coasts, which became known as *portolans*. Following the invention of the compass in the 12th century, portolans became more exact, and books containing sea courses and detailed information on distances and directions began to be produced. At some point, from these navigational notes emerged maps of the coasts with useful information – the 'portolan charts', which were the first nautical maps. These charts did not feature meridians or parallels. They were an attempt to describe the coasts and give details on harbours, geographical features and hazards for shipping. The toponymy (place names) was written perpendicularly to the coastline, and the inland areas were normally left blank except for decorative elements, though they did include drawings of compasses or wind roses from which emerged lines of bearings forming a web-like grid. The chart's scale was also given. Using a ruler, mariners drew a straight line joining the origin and destination points and it was translated in parallel fashion to a nearby wind-rose, which gave the bearing that had to be followed.

Although they achieved considerable precision, particularly on maps of the Mediterranean, the cartographic information was, obviously, very approximate. These portolan charts did not take the Earth's curvature – its roundness – into account, nor were they drawn up by using any mathematical projection whatsoever.



A map of Europe and the Mediterranean from the Catalan Atlas of 1375; the illustration is a copy from the 19th century.



Original Mercator map, 1569.

J.M. Sachs' article *A Curious Mixture of Maps, Dates and Names* (1987) points out that, although logarithms appear in the Mercator projection's mathematical equation, John Napier did not publish his work on logarithms until the beginning of the 17th century. What's more, to get these equations it is necessary to make use of the tools of calculus applied to differential geometry, but Newton and Leibniz were not born until 50 years after Mercator died, and Gauss did not devise differential geometry until the beginning of the 19th century. So, how did Mercator draw his map of 1569? Despite not having the tools that would eventually be developed in mathematics, Mercator had a profound scientific knowledge of the cartographic problem and consequently a very well-developed intuition. Mercator's work was practical and based on large data tables. However, the fact is that he did not leave any technical explanation of his map, or any nautical tables related to it, and not even a practical guide on how mariners should use it. That, and also because as far as sailors were concerned he was basically a scientist and not one of them, is why, perhaps, this map had little impact on the world of cartography for 30 years. There were some exceptions, such as that of his friend and fellow cartographer Abraham Ortelius, who included eight maps produced by Mercator's projection in his atlas entitled *Teatrum orbius terrarum* (1570).

The projection's mathematical explanations were to come from Cambridge mathematician Edward Wright (1561–1615) who, in his book *Certain Errors in Navigation, Detected and Corrected* (1599, and an extended version in 1610), in addition to including new nautical tables and giving instructions for obtaining fixed directions on maps based on the Mercator projection, provided mathematical explanations on how to generate the map, together with some intuitive descriptions. For example, he imagined the reference sphere as a ball, or a globe, which is inside a cylinder tangential to the sphere at the Equator and pumped up with air so that the globe is flattened against the cylinder and the image of each point on the reference sphere is the point of contact with the cylinder.

Initially, cartographers were slow in taking up this projection. The Dutch cartographer Peter Plancius used it in 1594 to draw up nautical charts and Jodocus Hondius used it for his world map in *Typus totus orbis terrarum* (1597), among others. The first maritime atlas ever was designed with the Mercator projection between 1646 and 1647 by the Englishman Robert Dudley.

TYPUS TOTIUS ORBIS TERRARUM, IN QUO
& Christiani militis certamen super terram in pietatis studiosi gratiam graphice designatur, à Iud. Hondio cælatore
NOVA FRANCIA
AMERICA AN DOM
TARTARIA
MAR DEL NORT
MAR DEL ZUR
PACIFICUM MARE
TERRA DEL FUEGO
PECCATUM
DIABOLUS
MUNDUS
MORS
CARO

A map entitled Typus totus orbis terrarum *(1597), also known as the 'Christian Knight Map', by Jodocus Hondius, made with a Mercator projection. In the map's central area, the Christian gentleman can be seen battling against Sin, the Flesh, the Devil and Death. Additionally, the World offers him the poisoned chalice of the Whore of Babylon, used at times as a symbol of the Roman Catholic Church.*

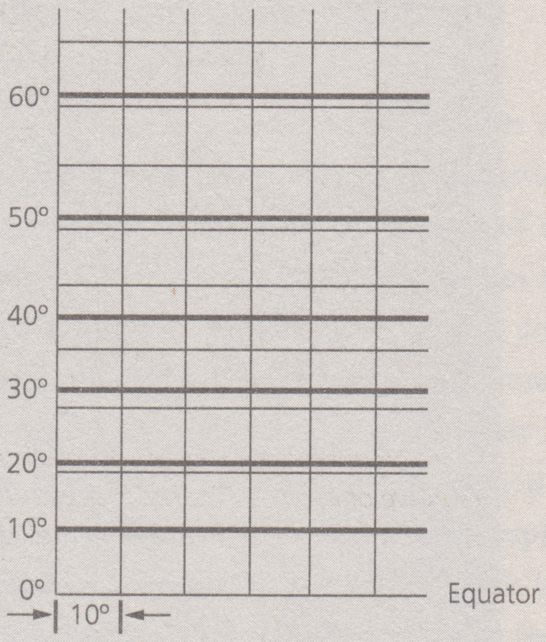
`

THE MERCATOR PROJECTION'S MATHEMATICS

To get an idea of the height, in respect to the Equator, at which the parallels on the Mercator map have to be placed, we can make small increases in the latitude on which to apply the corresponding factor of scale. If we start from a parallel of latitude ϕ and we take small intervals of size t, we get a series of latitudes marked, t, $2t$, …, $\phi - t$, ϕ, onto which we stretch out the height. As the distortion in the direction of the meridian for a latitude α has to be, from what we have seen, equal to the distortion in the direction of the parallel, which is $\sec\alpha$, then the distortion vertically in the marked latitudes will be $\sec t$, $\sec(2t)$, …, $\sec(\phi - t)$, $\sec\phi$, and as the length of the sphere between the latitudes marked is t, thus the calculation of the height of the parallel of latitude ϕ will be:

$$t \sec t + t \sec(2t) + \ldots + t \sec(\phi - t) + t \sec\phi.$$

For example, if we want to calculate the height of the parallel of latitude $\phi = 60°$ and we take intervals of $t = 10°$, as $\sec 10° = 1.0154$, $\sec 20° = 1.0642$, $\sec 30° = 1.1547$, $\sec 40° = 1.3055$, $\sec 50° = 1.5557$ and $\sec 60° = 2.0000$, by multiplying by 10 and adding we would get 80.955, in other words, parallel 60° should be placed at the height on which that of 80.955° would be if it were evenly spaced (as shown in the image).



That is what Edward Wright did, and – we imagine – Mercator did something similar. Now let's look at it in a more modern way. For a cylindrical projection, on which the Equator is the axis of the x and the parallel of latitude ϕ is a horizontal line at a height $y = h(\phi)$, as the scale factor (distortion) in the direction of the meridians λ has to be equal to the scale factor in the direction of the parallels $\mu = 1/\cos\phi = \sec\phi$, the result is:

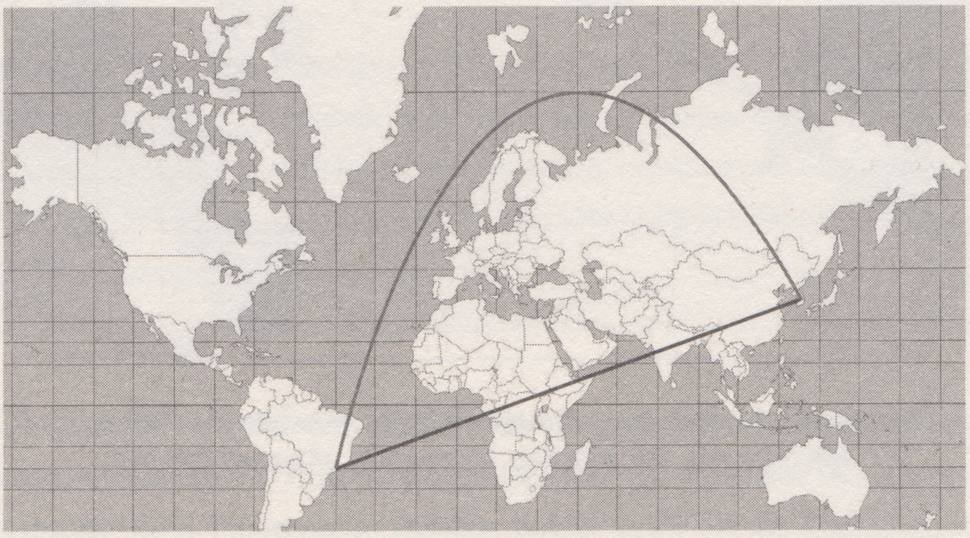
$$\lambda = \lim_{t \to 0} \frac{\Delta dist_{\text{map}}}{\Delta dist_{\text{sphere}}} = \lim_{t \to 0} \frac{h(\phi + t) - h(\phi)}{t} = h'(\phi) = \sec\phi$$

Then,

$$h(\phi) = \int_0^{\phi} \sec t \, dt = \ln\left[\tan(\phi) + \sec(\phi)\right] = \ln\left[\tan\left(\frac{\pi}{4} + \frac{\phi}{2}\right)\right].$$

But let's get back to Mercator's projection and recall the properties of a map produced with it:

1. It is a rectangular map, as it is made by a cylindrical projection.
2. The meridians and parallels form a grid of straight lines that intersect at right angles.
3. It is a conformal projection, on account of the construction itself, which does not preserve distances, areas, geodesics or the shapes of large areas.
4. The distortion of areas, shapes and distances is very little near the Equator (a line where the scale is true), but it increases greatly as it gets nearer the poles, which means it is very appropriate for maps of areas close to the Equator.
5. The loxodromics, or rhumb lines, transform into straight lines.

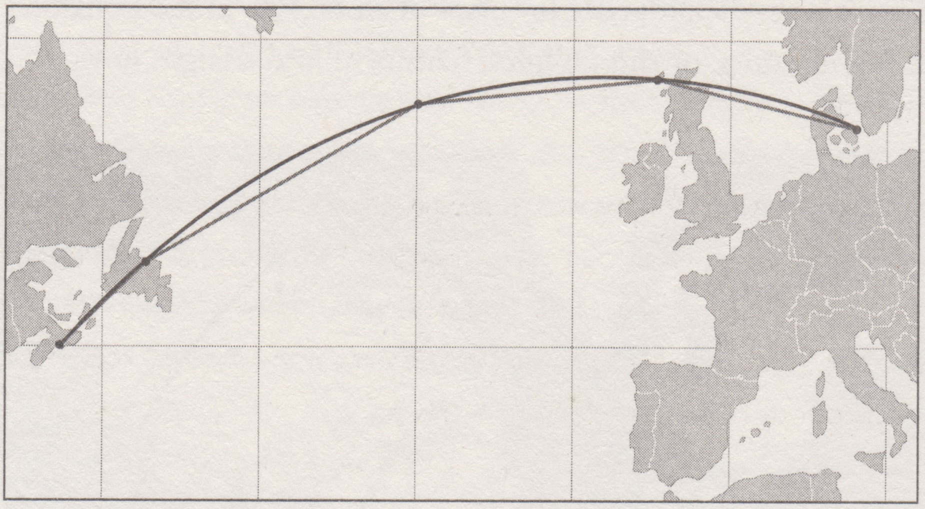


Comparison between the loxodrome, or rhumb line, and the orthodrome, the curve of minimal distance, between the cities of Rio de Janeiro and Seoul, on the Mercator map (source: Carlos A. Furuti).

This map was Mercator's dream come true. If a mariner wants to sail from point *A* to point *B* on the terrestrial surface, they only need to draw the straight line that joins both points on a Mercator-type nautical chart and to calculate the bearing marked by that straight line and head their ship in that direction. However, as we know, the loxodromes are not orthodromes, and although they show the path that is simplest to follow – as it is only necessary to maintain a constant direction – they do not give the shortest route between origin and destination. However, sailing on this latter path may turn out to be difficult as it requires repeated changes of direction.

Sailors and airplane pilots ended up by finding a compromise solution. To travel between the origin and the destination, they would do the following:

1. Draw a geodesic (straight) on a central map, or even on an equidistant azimuthal map centred on the origin.
2. Break up the geodesic into pieces, creating a set of strategic points.
3. Translate these points to a Mercator chart and join them by straight lines, which are the loxodromes, and mark the direction that should be taken on each of the strategic points.



Method of great circle approximation by loxodromes, used by mariners on a Mercator map and with help, for example, of a gnomonic map.

The Mercator projection has undoubtedly been, and still is, the projection par excellence for designing nautical and air navigation charts since its introduction in the 17th century. It has been the standard projection used by the US Coast and Geodetic Survey (now part of NOAA) since 1910, and by Spain's Instituto Hidrográfico de la Marina, to give just two examples.

Perhaps due to the importance of this projection during an epoch of voyaging, the truth is that the use of the Mercator projection became very common and was one of the cartographic transformations that was most used in world atlases up to the 20th century, despite the strong distortion that it generates in zones close to the poles. It was used for wall maps, in text books, in atlases in general, including school atlases, in popular publications and in the press. US cartographer John Snyder (1926–1997), of

CHARLES LINDBERGH'S FLIGHT

The American aviator Charles Lindbergh (1902–1974) became world famous for being the first aviator to make a solo, non-stop transatlantic flight. In 1919, a wealthy New Yorker offered a prize of $25,000 to the first pilot to fly solo non-stop between New York and Paris. Lindbergh believed he could win the prize with the correct plane and convinced a group of businessmen in St. Louis to finance the exploit, which included building a special plane that he had helped to design, named 'The Spirit of St. Louis'. On 27 May 1927 he took off from Long Island with 4 sandwiches, 2 water bottles and 1,700 litres of petrol; 33.5 hours and 3,610 miles (approx. 5,800 km) later, he landed in Paris in front of a crowd of 100,000 spectators, who were there waiting for him. The 'lone eagle', as they nicknamed him, had become a world hero. Lindbergh planned his journey thoroughly with navigation charts. In his own words:

"…I spent the greater part of the construction period working out the details of navigation and plotting the course, with its headings and variations, on the maps and charts. After working out the track on the gnomonic and Mercator charts, I checked over the entire distance from New York to Paris with the nautical tables... From New York to Paris I worked out a great circle, changing course every one hundred miles...".

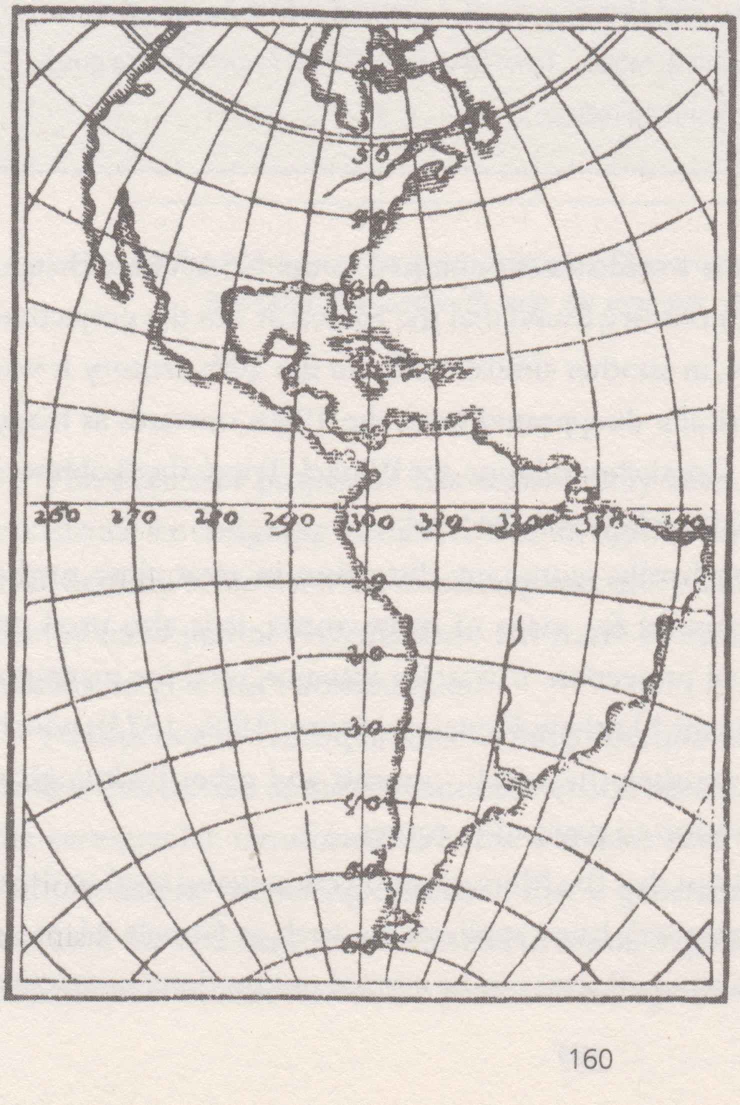
the USGS, in a study on different world atlases published in the USA, Great Britain France and Germany in the 19th century, found that the Mercator was the projection most commonly used. However, in another similar study on the 20th century it was found that the projection practically disappeared from the 1940s onwards as many other different kinds such as the Goode homolosine, the Winkel-Tripel, the Robinson, the Eckert IV or the Van der Gritten began to be used.

The fact that the projection hardly causes any distortion in areas close to the Equator means that it is very useful for maps of those zones. It is also used on account of its being a conformal projection. It was, for instance, used for maritime maps made by US Navy lieutenant Matthew Fontaine Maury (1806–1873), which included information such as weather, the winds, currents and other hydrological and meteorological data, in addition to navigation routes.

To conclude, we must point out that the Mercator projection is also used to draw world maps for some of the Internet's latest applications, such as Google Maps or Virtual Earth. Shifts and zooms on small areas can be carried out on these interactive maps with only a small amount of distortion as Mercator's projection is conformal, which means that shape distortion is minimal at a local level and for small areas.

The transverse Mercator projection

The Mercator projection which we have just studied is a conformal transformation on which there is little distortion in areas close to the Equator, which is the curve of tangency of the cylinder with the reference sphere. Consequently, if we take the transverse aspect of this projection, in other words, if we keep the same cartographic transformation but now the cylinder to which we project the reference sphere is turned 90°, so that now the line of tangency is a meridian, we will have the transverse Mercator projection, centred on that meridian. This is also conformal and produces low distortion at the zones close to the meridian where the tangency is set. All these factors make it very appropriate for maps of land zones spreading north-south, such as the American continent or India.



A map of America drawn up with the transversal Mercator projection, which is the image that Lambert showed as an example of his projection.

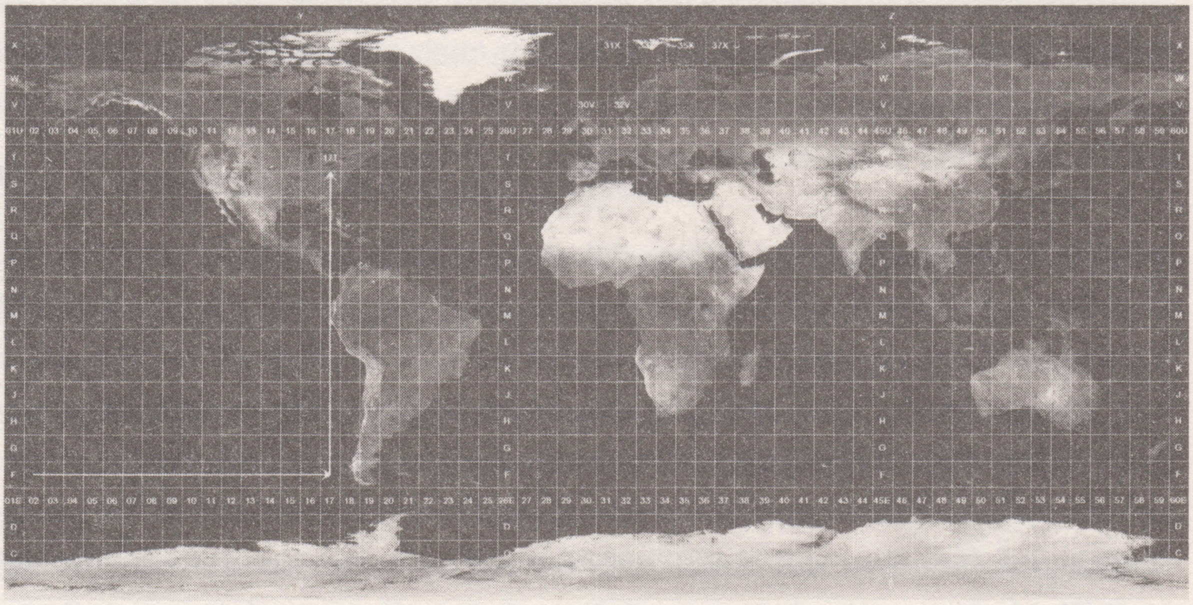
This cartographic projection was one of those introduced by Lambert in his work in 1772, and was later studied in 1822 in their ellipsoidal version by Gauss and the mathematician and topographer Louis Krüger (1857-1923). It is therefore also known as the Gauss-Krüger projection. Its properties are:

1. The meridians, parallels and, in general, the loxodromes, are not straight lines.
2. It is a conformal projection and so preserves the angles and the shapes at local level.
3. The distortion in zones near the central meridian is very small (in fact, the scale is true on that meridian) and increases as we move away from it.

This transformation is therefore ideal for making both maps extending in the north-south direction and for maps of zones that are not very large simply by choosing a suitable central meridian that passes through that area. It has been used in particular in different atlases for drawing up maps of North America, the western region of the former Soviet Union, India, the Far East, South East Asia, the eastern area of Australia and Africa. In general terms, it is widely used, for example, for nearly all European countries. But, basically, the advantages of this projection for smaller areas meant that it has been used as the basis of coordinate systems for topological maps, such as, for example, the British grid system of coordinates (in 1919), or the SPCS, a system of projections for the USA (in 1930). However, the definitive system, and today we can say the universal one, was developed in 1947 by the U.S. Army Map Service – the UTM projection service (Universal Transverse Mercator).

The Earth between latitude 84° N and 80° S is divided into 60 zones of 6° of longitude, and for each of them the transverse Mercator projection is used, with the meridian of tangency being that situated in the middle of the area concerned. The zones are numbered from 1 to 60. Meanwhile, letters are designated in accordance with a division made of 20 zones of 8° of latitude, from south to north. Thus, Bilbao is in UTM zone 30T, New York in 18T, Sydney in 56H, and Alexandria in 35R. For the polar zones, outside latitudes 84° N and 80° S, the UPS (Universal Polar Stereographic) system is used.

The UTM system is the one universally used for maps with scales less than or equal to 1:500,000 by the majority of topographic, geological, geodesic, cartographic, military, maritime and similar agencies throughout the world. One of the world's most important agencies, the U.S. Geological Survey (USGS), has been using it since 1977.



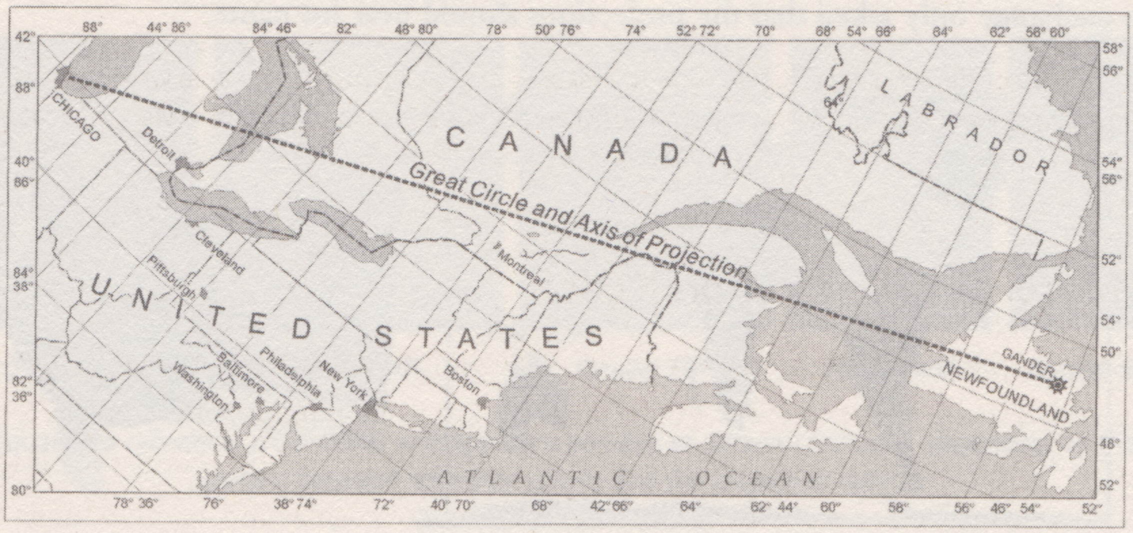
A map of the UTM zones. If we want to make a topographic map of the area where we are situated, we have to look and see what UTM zone we are in so as to find out which Mercator projection we should use.

The oblique Mercator projection

The oblique Mercator projection might be taken into consideration, that is, the case in which the cylinder of projection is tangential to the reference sphere on any great circle which is neither the Equator nor any of the meridians. This is how the oblique Mercator projection is obtained; this is obviously conformal and again exhibits the important property of producing little distortion in the areas around the great circle of tangency. That property makes it suitable for areas extending around a great circle.

The origin of this projection is not very clear, but among its first maps were one of Switzerland in 1903 by Max Rosenmund, and a map of Madagascar in 1928 by Jean Leborde. Since then it has been used, for example, to make maps of the American continent (both as a whole and for partial zones), Eurasia, Australasia and for smaller territories such as the West Indies (the Antilles and the Bahamas), Hawaii, New Zealand, Italy and Alaska by the National Geographic Society and other agencies.

Little by little, throughout the 20th century, commercial transport was being developed, both for goods and for passengers. The emerging companies began to cover great routes, including transoceanic flights. Attempts were made to route these long-distance journeys along great circles so as to minimise flying time and fuel consumption. Aeronautical charts were usually long foldable maps oriented in the direction that joined the two airports and showed narrow strips along the route – basically they were not very useful from a cartographic point of view. On account of its properties, it was precisely the oblique Mercator projection that would be ideal for mapping air routes along great circles. So, in 1947, the U.S. Coast and Geodetic Survey used it for the first aeronautical chart for a route around a great circle, the one between Chicago and Gander. At that time, Gander airport, in Newfoundland, was an obligatory refueling stop for all transatlantic flights. This chart, as well as being centred on the orthodromic route, could be used to draw up other routes or measure distances around the former one as there was little distortion both in angles and distances.


CANADA
LABRADOR
UNITED STATES
ATLANTIC OCEAN
Great Circle and Axis of Projection
CHICAGO
Detroit
Cleveland
Pittsburgh
Montreal
Philadelphia
Baltimore
Washington
New York
Boston
GANDER
NEWFOUNDLAND
42°
88°
44° 86°
84° 46°
82°
48° 80°
78°
50° 76°
74°
52° 72°
70°
68°
54° 66°
64°
62°
58° 60°
58°
56°
88°
40°
86°
54°
56°
52°
84°
38°
54°
50°
82°
36°
48°
80°
52°
78° 36°
76°
38° 74°
72°
40° 70°
68°
42° 66°
64°
62° 44°
60°
58°
56°
46° 54°
52°

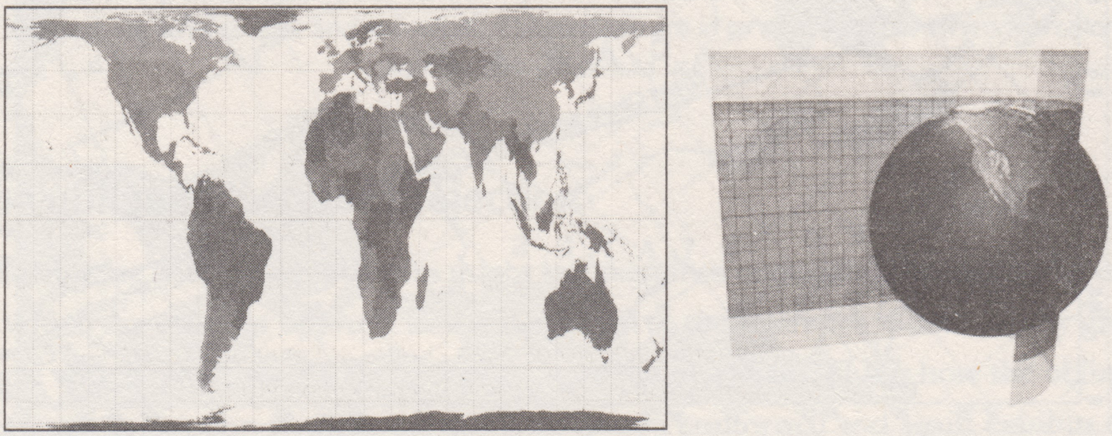
Aeronautical chart for the route Chicago–Gander made with the oblique Mercator projection.

The launching of space satellites, which NASA began in 1972, was to bring with it new cartographic applications for this projection. The satellites orbiting above the Earth almost along great circles (but not quite) took images of the surface. These images were not like the typical aerial photos, but consisted of a complex scanning of the terrestrial surface which generated a great deal of information. And that information had to be converted into planar images, on maps, with the least possible distortion. In 1976, John P. Snyder, Alden P. Colvocoresses and John L. Junkins, of the USGS, developed the Space Oblique Mercator projection from the original projection in order to solve the problem of how to show the satellite images.

Peters versus Mercator

We shall conclude this book by telling a story about manipulation in cartography which began in 1967 when the German historian Arno Peters presented his 'new' projection at a congress held at the Hungarian Academy of Sciences, although it would have no repercussions until 1973, when he presented it to the media in Bonn. But first let's take a look at the projection itself.

The Scottish clergyman James Gall (1808–1895) described a projection which was identical to Peters' projection and known as Gall's orthographic projection at a congress back in 1855, and he published it, together with two other of his cartographic projections, in the Scottish Geographical Magazine in 1885. He altruistically donated his three projections to whomever wanted to use them, and in exchange only asked that the projections should bear his name.

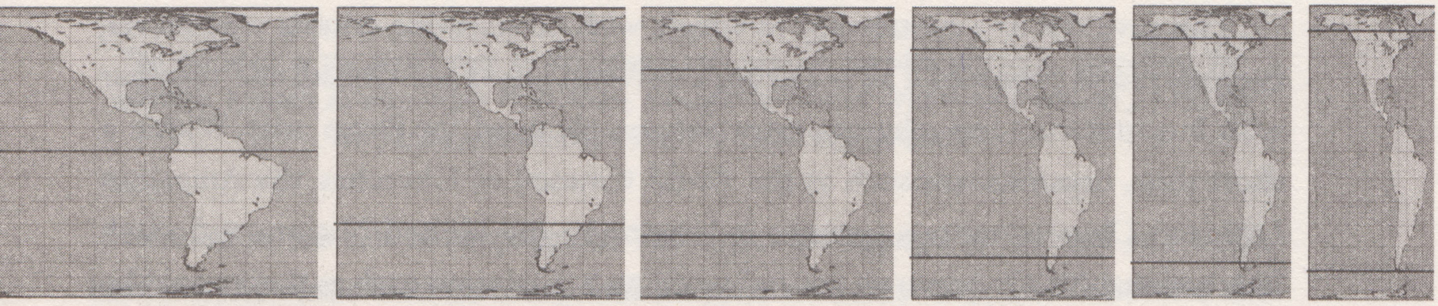


A map produced with the Gall, or Gall–Peters, orthographic projection, and a picture of the cylinder onto which the image is projected, secant to the sphere.

The construction of this projection is similar to the Lambert equal-area cylindrical projection which we saw in Chapter 5, with the difference that, instead of having a cylinder tangential to the reference sphere at the Equator (in its normal aspect), it uses a cylinder secant to the reference sphere which cuts the sphere at two parallels. The Gall orthogonal projection, later to be called Gall-Peters, takes the intersection parallels, which are to be standard curves on the map, as situated at a latitude of 45° from the Equator.The projection continues to be equal-area, the same as other similar ones that have been constructed with other standard parallels, such as W. Behrmann's in 1910, at 30°, those ofTrystan Edwards' in 1953, at 37° and 52°, or Hobo-Dyer's in 2002, at 37°, whose map became famous for having the south at the top.

ARNO PETERS (1916–2002)

According to his obituary, Peters was born in Berlin in 1916 and studied history, history of art and journalism at Berlin University. He was a film director in the agitated Germany of the 1930s, was awarded the title of 'doctor' in 1945 for a thesis on political propaganda, *The use of cinema as a means of propaganda*, and also worked as a journalist. He managed to go down in history as the creator of a solidarity and justice map, the famous Peters map, and expressed his thoughts on it in his book *New Cartography*. His main non-cartographic work was *Synchronoptic World History*, in which he carries out a revision of history and dedicates the same amount of space in the book to each century. In 1974 he co-founded the Institute of Universal History in Bremen.



Parallel 0° *Parallel 30°* *Parallel 37°* *Parallel 45°* *Parallel 50°* *Parallel 55°*

Equal-area cylindrical projections, with standard parallels at different latitudes, depending on the contact between reference sphere and the projection cylinder (source: Carlos A. Furuti).

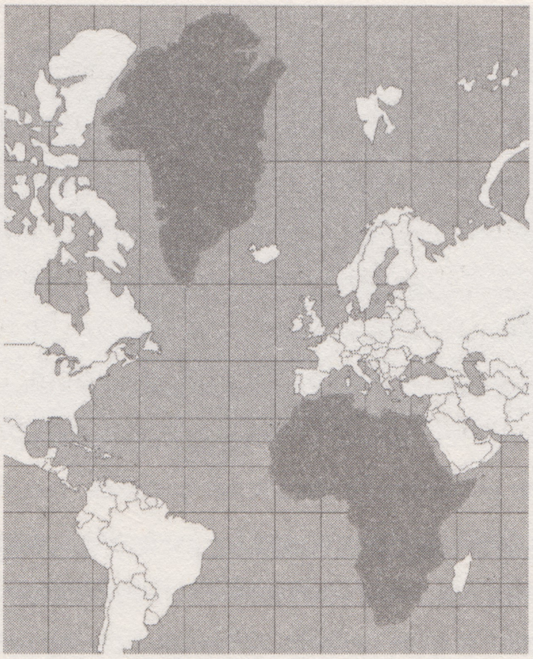
Arno Peters presented this projection to the cartographic world as original, and, when he was reminded that it had already been developed by Gall a hundred years earlier, he still defended it as a new and independent creation. The scientific community was reluctant to accept this map as a Peters' creation, though not for personal reasons but because in science value is attached to new, innovative creations and discoveries, whether they are in themselves original (a theorem, a theory or a map) or they provide an innovative proof or development. (For instance, it is common for some results or mathematical theorems to be proven for a second time, bringing simplicity and understanding to the original proof, or new techniques of proof.) But Peters' projection, the map-making community decided, did not fit the bill.

The second and most important part of the tale begins in 1973, when Arno Peters called a press meeting in Bonn. The journalists were handed a copy of his world map and a leaflet: "The Eurocentric nature of our geographic image of the world and its conquests". Peters presented his map as the only possible alternative, both from a social justice viewpoint and from the cartographic point of view, to the 'racist' and inadequate Mercator map. Peters' main argument was that Mercator's projection distorted areas in different parts of the world by showing the nations in the so-called 'third world' (Africa, South Asia and the central and southern part of America) as being small in comparison to the so-called 'first world' (North America, Europe and Russia). Peters protested that the former regions are inhabited by people with dark skin and the latter by people of lighter skin for that reason the Mercator projection was racist and its use had to be discouraged. After his attack on the Mercator map, Peters presented 'his' map of the world as the only possible alternative.

This gave rise to a controversy in which, with no scientific criteria whatsoever, the mass media presented it all as a struggle between David and Goliath and, together with humanitarian and religious organisations, plumped for Peters' map. Within a short time, organisations such as the World Council of Churches, the American Lutheran Church, various agencies belonging to the United Nations and some international NGOs began to use and promote it. Opinions such as these could be heard:

> "The Mercator map overvalues the white man and distorts the image of the world to the advantage of colonialists." (Peters)
>
> "The map of the future, for a world with more solidarity."
>
> "Ever committed to the cause of justice, he [Peters] chose the path of cartography to restore an image of the world in which all peoples would have the place in the world that was rightfully theirs, both geographically and politically."
>
> "Peters' map corrects the errors in Mercator's [...] it is more rigorous from a scientific point of view."

Peters took advantage of people's good faith and sense of solidarity, as well as their complete ignorance of map-making, to get his map acknowledged as the only 'solidarity map', and even worse, from a mathematical and cartographic viewpoint, presented it as the only 'accurate map'.



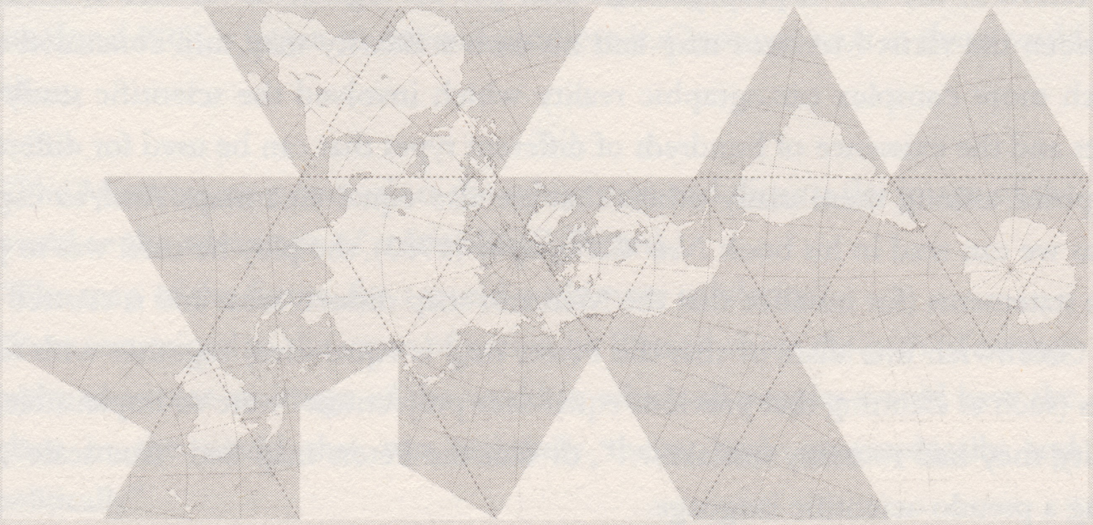
The distortion in the zones close to the poles is very high on the Mercator map. Greenland, for example, appears as even larger than Africa, though the area of Greenland is around 2,175,000 km^2, compared to the African continent's 29,800,000 km^2.

Versed in the arts of propaganda, Peters put his arguments forward as a battle between the racist Mercator map and his own solidarity map, and concealed the much more complex cartographic reality which involved the scientific study of maps and the existence of hundreds of different types that can be used for different purposes, among them many others that are also equal-area maps. And, to cap it all, as we can read in his book *New Cartography* (1983), the ploy he used was to put true statements (for instance that the Mercator map distorts areas and is centred on the Greenwich meridian, or that the Peters map is equal-area) together with false ones (such as claiming that previous equal-area projections "were so implausible, in reality, they had so many mistakes…", or that the Peters map had "true scale", all using a pseudo-scientific language.

At that time, society was already progressing, though maybe too slowly, towards the use of world maps different from Mercator's, precisely because cartographers were aware that it was an excellent projection but not for showing the world as a whole, on account of the distortion it produced in area. For example, in Snyder's study on the projections used in atlases in the 20th century it was shown that the Mercator projection was already hardly being used, and that others were being adopted. What Peters did, however, was to get the decadent hegemony of Mercator's world maps replaced by his own empire, thrusting aside all the rich abundance of

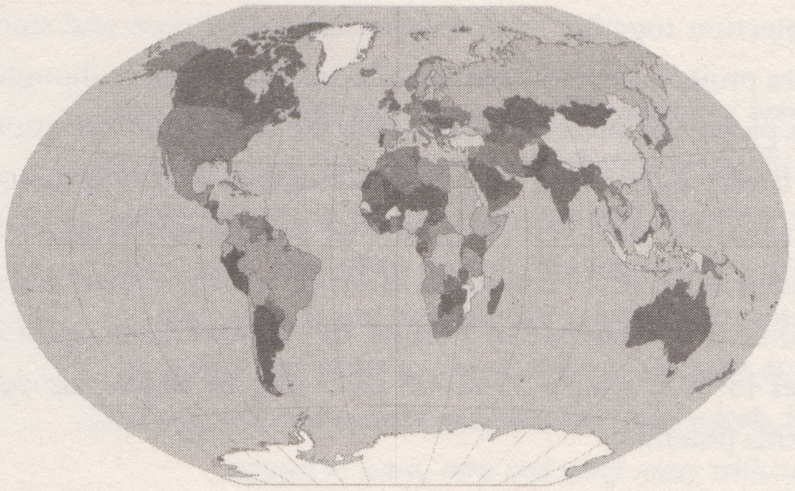
THE DYMAXION MAP

Buckminster Fuller, creator of the geodesic dome, designed a map that was an attempt to somehow offer a planar image of the terrestrial surface that would be more universal than that of other world maps. His idea was to project the terrestrial surface onto a regular, or semi-regular, polyhedron, and afterwards to open it out on the plane. In his projection named Dymaxion (DYnamics MAXimum tensION; not a trade mark but rather the philosophy behind his designs and pieces), patented in 1946, he used a cube-octahedron (8 triangular sides and 6 square ones), while in his 1954 version he used an icosahedron (20 triangular sides) with a slight modification. The projection he used is not gnomonic but, instead, is produced by means of a construction similar to the one used for the geodesic dome. The map generated by the Dymaxion projection has only a small amount of distortion in areas and shapes and, furthermore, that distortion is quite uniform. Although it can be unfolded in several ways, the most common way to do it is with the North Pole more or less in the centre. It provides a vision of the world in which north and south no longer exist, but rather it can be shown with any orientation. The continents are not separated and appear to be more like an island surrounded by the ocean.



An unfolded Dymaxion map based on an icosahedron. The lines for folding are marked on it (source: Eric Gaba).

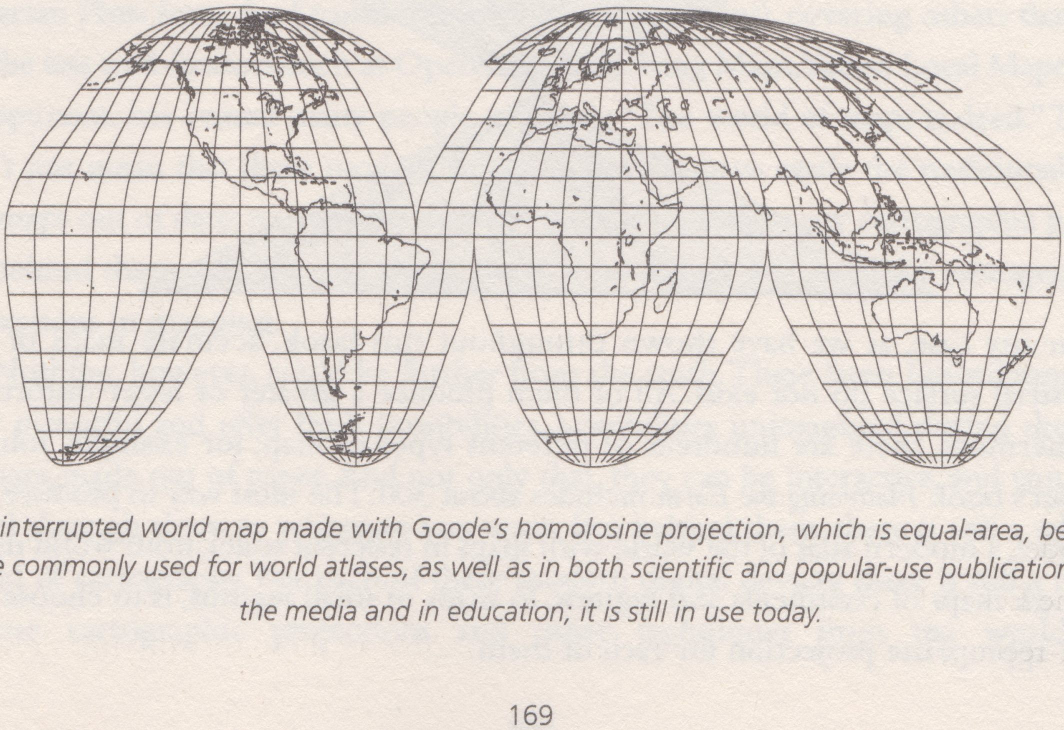
cartography both in world maps that preserved the area (such as those produced by Goode's homolosine projections, the Mollweide, the Sanson-Flamsteed sinusoidal or the Eckert IV), and in those that preserved other properties (for example, Miller's equi-rectangular) and other compromise maps (that is, those that did not preserve



The Winkel–Tripel projection is a compromise, in other words it preserves no metric properties, but the distortion in each of them is small.

any specific property but whose distortion was very slight, in particular in regard to shape and area, such as those used by the National Geographic Society, the Arthur H. Robinson projection, from 1961, and the Winkel-Tripel, from 1921). All are examples of maps of the world that were becoming common in the second half of the 20th century.

The indignation of the cartographic community was due, in part, to society's disrespect for their science and their work, as well as the manipulative arguments used to back Peters' world map. To start with, the existence of maps that preserved area originates back in Lambert's 1772 article, in which he presented his equal-area

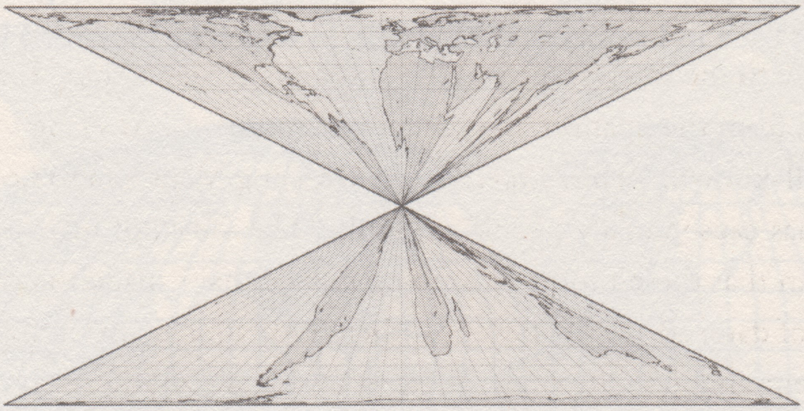


The interrupted world map made with Goode's homolosine projection, which is equal-area, began to be commonly used for world atlases, as well as in both scientific and popular-use publications, in the media and in education; it is still in use today.

cylindrical projection together with another azimuthal type, and since then many other equal-area projections had been introduced. What's more, though it is true that the Gall-Peters projection preserves areas, it has the defect that distortion in the shape of the regions is very high, being very much elongated in the north-south direction in the central zone, and then shrinking beyond parallels 45° N and S. Ironically, the distortion in shape is heavily accentuated in Africa and the central and southern part of America, while it is moderate in Europe, the United States and Canada, which are near parallel 45° N. A couple of quotes and a map by Snyder are very interesting in relation to this polemic:

> "[the Peters world map] isn't any better than similar maps that have been in use for 400 years."
>
> "The promotion of the Peters' projection thus seems to have taken on some of the same single-mindedness as claims for 'one true faith' or a patented panacea. In the process of trying to raise social consciousness with his map, Peters has cast aside objectivity and important facts."

Snyder's map shown here was his way of expressing, with humour and irony, that it is not enough for a map to be equal-area, but that other issues have to be taken into account, and one of them is the property of preserving the shapes of continents and countries.



In any case, as we have shown throughout this book, accurate maps of the terrestrial surface do not exist. All of them produce a greater or lesser distortion. Furthermore, there are hundreds of different types of map; for example, John P. Snyder's book *Flattening the Earth* includes about 300. The ideal way to produce, for instance, a modern atlas of the world with maps in different scales, from world maps, through maps of continents and regions, to maps of small regions, is to choose the most appropriate projection for each of them.

Epilogue

> *Once upon a time there was a map. The map had been used for years and years. A long time. It gave routes, showed where the paths ran, and helped folk not to get lost. That's the task that every map must carry out – to be there when needed. And this one was, without doubt. In recent times it had been perused so often, that someone must have thought it was better to spread it out on the table and leave it there. That way, whoever wanted to could come, take a look, learn what they wanted to learn, turn and walk away without wasting time. It was a map, a good map.*
>
> *But one day it stopped being useful. Who knows why, but it stopped providing the service it had so faithfully given the community to which it belonged. Perhaps because it grew old. Perhaps because what it depicted no longer corresponded with reality.*
>
> Alber Vázquez, *Instrucciones para doblar un mapa* (2004)

The technological progress that has been made in recent years in the world of cartography with images provided by satellite, GPS systems and the great number of other tools that have flourished on the Internet, from the powerful Google Maps and Street View to the interesting SIGPAC (System of Geographic Information on Agrarian Plots from the Spanish ministry of environment), covering others that are maybe less well-known, such as OpenStreetMap, Bing Maps, Yahoo Local Maps and Mappy.com, has caused many people to think: "The world of maps is dead." They don't just mean that these modern mapping devices have made the traditional paper maps out of date; no, they believe that there is no longer any cartographic process behind them and, for one thing, there is no longer any need for cartographic projections in mapping.

Nothing, however, could be further from the truth. These latest innovations are very powerful and offer fresh possibilities, possibilities unimagined during the era of maps made out of paper. And not only that, they can be interactive, and you can travel through them to wherever you want. But behind them, there is still a whole world of cartography. For each of these modern digital devices there is still a need to use cartographic projections and other techniques from the world of cartography. Thus, for example, the satellite images we normally see are not 'photographs' taken from space; but rather, they are the product of following process. Firstly, the satellites take a scan of the terrestrial surface, generating a huge amount of information. This data is then transformed into an image by means of a cartographic projection, whether it be the Space Oblique Mercator or any other of the ones we have seen throughout this book. Google Maps uses the Mercator projection which, being conformal, preserves the shapes at local level, which is very handy for interactive maps to be used for zooming. GPS satnavs show maps drawn with the UTM (Universal Transverse Mercator) system or some other cartographic projection.

In short, the new digital tools are certainly a new and potent means of transmitting cartographic information, but the science of cartography has not disappeared – it still carries on the same. Eye-catching they may be, but these new cartographic tools are no more than the latest step in the long history of cartography.

Bibliography

DILLER, A., 'The Ancient Measurements of the Earth', *Isis*, vol. 40, no. 1 (Feb. 1949), pages 6-9.

FEEMAN, T.G., *Portraits of the Earth, A Mathematician Looks at Maps,* AMS, Providence, 2002.

FURUTI, C., *Map Projections*: http://www.progonos.com/furuti/MapProj/CartIndex/cartIndex.html

MONMONIER, M., *Rhumb Lines and Map Wars,* Chicago, The University of Chicago Press, 2004.

POLKING, J.C., *Mapping the Sphere,* Houston, Rice University: http://math.rice.edu/~polking/cartography/

RICKEY, V.F., 'How Columbus Encountered America', Washington, *Mathematics Magazine*, 65, no. 4 (Oct. 1992), pages 219-225.

ROBINSON, A.H. ET AL., *Elements of Cartography,* New Jersey, John Wiley and Sons, 1953 (6th edition, 1995).

SNYDER, J.P., *Map Projections, A Working Manual*, Reston, USGS Professional Paper 1395, 1987.

—: *Flattening the Earth, Two Thousand Years of Map Projections*, Chicago, The University of Chicago Press, 1993.

SOBEL, D., *Longitude*, London, Harper Perennial, 1997.

Index

Anaximander 14
angle 26-42, 62-75, 87-1025, 107-124, 129-134
Archimedes 25-26, 71-75
area 64-69, 74-76, 82-85, 129-133
Aristotle 14, 15, 16, 25, 88
Arzachel 117-118
atlas 81, 100-102, 104, 150, 155
azimuth 96-97, 100

Bible 20-21
Blaeu, Joan 118

Carroll, Lewis 129
cartogram 84
Cassini, Jean-Dominique 22, 33-34, 43, 113
chart
cholera, map of 53
Columbus, Christopher 20, 31-32, 42
cone 68-69, 95, 123-128, 135
conformal, projection 67-69, 123, 151
cosmological myth 15
curvature
  geodesics 50
  mean 136-138, 140
  of Gauss 135-140
  of the circumference 49-50
cylinder 68-79, 95, 135-139, 155, 164

direction 54, 62, 102, 108, 151, 153, 157
distance 26-31, 46-49, 57-66, 99-102
distortion
  in the direction of the meridians 75, 91, 156
  in the direction of the parallels 75, 91

eclipse 16, 37, 43, 88, 118
ecliptic, plane of the 38-39, 114
ecumene 12, 14, 141
ellipsoid 13, 22-24, 30, 36, 56
equator 22-23, 77-80, 90, 155-160
Eratosthenes 26-30, 37, 41, 78, 145
Eudoxus 25-26
Euler, Leonhard 134-135

Finé, Oronce 126, 150
fish-eye (photographic lens) 121
Flamsteed, John 104
formula
  of Gauss-Bonnet 140
  of the sum of the angles of a triangle spherical 133, 139, 140
  of the sum of the angles of a planar triangle 140
Fuller, Buckminster 48, 103, 168

Galilei, Galileo 43, 104
Gall, James 164-165
Gauss, Carl Friedrich 135-140, 154, 160
geodesic 11, 22, 24, 52, 122, 143
geodesic domes 48-49
geodesics 36-37, 46-50, 64-69, 129-132
geographical projections 13, 35, 37, 45
geoid 22-24
geometry 11, 46, 72, 74, 109, 121, 140
  differential 10, 52, 58, 62, 121, 139, 154
globe
  of reference 56-57, 103
  terrestrial 16, 25, 36, 140-144, 150
gnomon 27-30, 39, 87
Google Maps 160, 171, 172
Goos, Peter 118
GPS 23, 38, 717, 172
gravity 11, 22-24
great circle 35-50, 57-64, 92, 102, 162-163

Halley, Edmond 43, 111, 113
Harrison, John 44-45
Hipparcus 12, 37, 43, 105
homothecy 57, 74
Hondius, Joducus 56, 106, 119, 120, 126, 155

Indicopleustes, Cosmas 20-21
isometry 58, 63, 64, 129-134, 139

Kepler, Johannes 102, 119

Lambert, Johann Heinrich 71-73, 123, 127, 160, 169
latitude 22-23, 35-40, 90, 156, 161, 165
law of universal gravitation 22, 33
Library of Alexandria 26-27, 30, 120
Lindbergh, Charles 159
longitude 25-29, 33-34, 40-45, 152
loxodromes 102, 108, 122, 151, 157-158
L'Isle, Guillaume de 114

map
  celestial 115
  Dymaxion 7, 168
  fantasy 55
  National Topographic (Map) of Spain 161
  polyhedric 103
  topological 55
Maslama 105, 106
mathematical cartography 51-52, 57, 139
Mercator, Gerardus 81, 118, 141, 146, 149-156
meridian 26-30, 35-50, 73-81, 123-125
  of Greenwich 41-42, 85, 167
metre 30, 33, 34
Morgan, Augustus de 61, 87, 101

National Geographic Society 81, 83, 146, 168
Nazi propaganda 86
Newton, Isaac 22, 33, 43, 44, 154

Ortelius, Abraham 151, 154
orthodrome 87, 101, 102, 122, 151, 157, 163
orthophotography 70

parallel 35-40, 74-81, 124, 156, 164-165
Peters, Arno 164-170
plane tangent to a surface 136
planispherium (planisphere) 12, 61, 105, 106
Plato 15, 26
polyhedron 103, 168
portolan, charts 149, 151, 153
portolans 153
Posidonius 31-32
principal curvatures 138
projection
  algorithmic 66, 68-69, 78, 96, 123
  armadillo (or of Raisz) 96
  azimuthal 87, 97, 105, 108
    equal-area of Lambert 71, 85, 99
    equidistant 98, 99, 145-148, 158
  cartographic 10, 54-58, 67, 172
  conic 123-126, 128, 148
    bipolar oblique conformal 10, 123
    conformal of Lambert 10, 71, 123, 127-128
    equal-area of Albers 10, 68-69, 85, 124, 125, 128
    equidistant 128, 148
  Craig (or Mecca) retroazimuthal 98
  cylindrical 71, 74-83, 156, 157, 165, 169

    equal-area of Lambert (or of Archimedes) 71-87, 164
    of Miller 79
Eckert IV 10, 85, 159, 167
equal-area 65, 67-86, 97, 99, 130-131
    of Behrmann 10, 77, 78
equidistant 97, 120, 144-148
equirectangular (or rectangular) 144, 145, 168
Gall's orthographic (or Gall-Peters') 10, 77, 85, 164
geometric 68, 76, 78, 90
gnomonic (or central) 87, 90-94, 98-103
Goode homolosine (interrupted) 159, 167, 169
Hammer-Aitoff 69
Hobo-Dyer 164
in general perspective 66-67, 70
Lagrange 71, 123
Mercator 8, 85, 112, 149-170, 172
    oblique 162-163, 172
    transversal 116, 149, 160-162, 172
Mollweide 10, 69, 81, 85
octahedron projection (or Cahill's Butterfly) 103
of Braun 77-79
of Gall 77-79
plate carrée (or planar chart) 78, 144-145
polyconic 125
pseudoconic 125, 126
pseudocylindrical 80-81, 85
Robinson 10, 69, 81, 83, 159, 168
sinusoidal, or Sanson-Flamsteed 69, 81, 85, 142, 150, 167
stereographic 69, 77, 78, 105-128
Trystan Edwards 164
Van der Gritten 10, 146, 159
Winkel-Tripel 10, 69, 83, 159, 168-169
Ptolomy, Claudius 11-12, 22, 32, 105, 125, 126
Pythagoras 15

radius of the Earth 56

scale 9, 23, 54-58, 82-83, 127, 144
scenographic, projection 66
Snyder, John 112, 120, 159, 163, 167, 170
Society for the Diffusion of Useful Knowledge, SDUK 100
Speed, John 121, 126
sphere 23, 46-50, 57-79, 107-108, 142
spherical gores 141, 142
spherical triangle 131-133, 139
standard curve (or line) 123, 144, 145, 148, 164

Thales of Miletus 14, 87, 88
theorem
    of Thales 88, 93, 100, 110
    of the four colours 61
Tissot indicatrix 76, 92
Toscanelli, Paolo 32
triangulation 32-34

UPS (Universal Polar Stereographic) 116, 161
USGS (United States Geological Service) 116, 121, 127, 145, 159, 161, 163
UTM (Universal Transverse Mercator) 116, 149, 161-162, 172
vector normal to the surface 136

Waldseemuller, Martin 141-142
Werner, Johannes 43, 120, 126
Wright, Edward 155, 156

Ways of Mapping the World

Maths and cartography

Many of us recall world maps, displayed on the classroom wall, in which a gigantic Greenland existed alongside a tiny South America. Why were the dimensions of the continents so obviously wrong? The reason, we now know, is that when attempting to project the surface of the globe on to a two-dimensional plane, we have no choice other than to distort reality. The perfect flat map belongs in the same category as unicorns or proofs of squaring the circle: it does not exist! However, the history of how this conclusion was reached and the attempts to solve this problem constitute an absorbing episode in the history of mathematics.